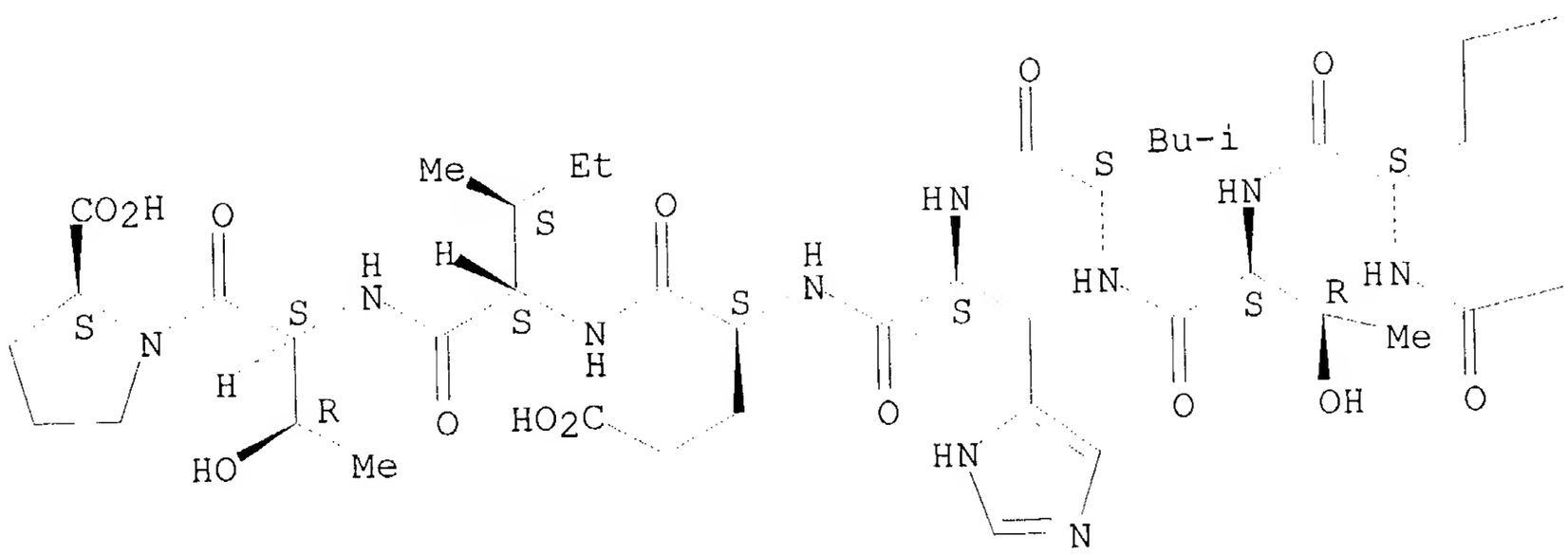
```
=> fil reg
FILE 'REGISTRY' ENTERED AT 16:23:16 ON 14 JUN 2004
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2004 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES:   13 JUN 2004  HIGHEST RN 692726-52-6
DICTIONARY FILE UPDATES:  13 JUN 2004  HIGHEST RN 692726-52-6

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2004

  Please note that search-term pricing does apply when
  conducting SmartSELECT searches.

Crossover limits have been increased.  See HELP CROSSOVER for details.

Experimental and calculated property data are now available. For more
information enter HELP PROP at an arrow prompt in the file or refer
to the file summary sheet on the web at:
http://www.cas.org/ONLINE/DBSS/registryss.html

=> d que l1
L1             1 SEA FILE=REGISTRY GGG['GLU''GLA']VR['GLU''GLA']SA['GLU''GLA']TL
                 H['GLU''GLA']IT[P'HYP']|GGG['GLU''GLA']VR['GLU''GLA']SA['GLU''G
                 LA']TLH['GLU''GLA']ITP/SQSP


=> d sqide l1 1

L1   ANSWER 1 OF 1  REGISTRY  COPYRIGHT 2004 ACS on STN
RN   459871-55-7  REGISTRY
CN   L-Proline, glycylglycylglycyl-L-α-glutamyl-L-valyl-L-arginyl-L-
     α-glutamyl-L-seryl-L-alanyl-L-α-glutamyl-L-threonyl-L-leucyl-L-
     histidyl-L-α-glutamyl-L-isoleucyl-L-threonyl- (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   114: PN: WO02072005 SEQID: 170 unclaimed sequence
FS   PROTEIN SEQUENCE; STEREOSEARCH
SQL  17

PATENT ANNOTATIONS (PNTE):
Sequence |Patent
Source   |Reference
=========+===========
Not Given|WO2002072005
         |unclaimed
         |SEQID 170


SEQ      1 GGGEVRESAE TLHEITP
           ========== =======

HITS AT:   1-17
MF   C74 H120 N22 O29
SR   CA
LC   STN Files:   CA, CAPLUS, TOXCENTER, USPATFULL
DT.CA  CAplus document type:  Patent
RL.P   Roles from patents:  PRP (Properties)
```

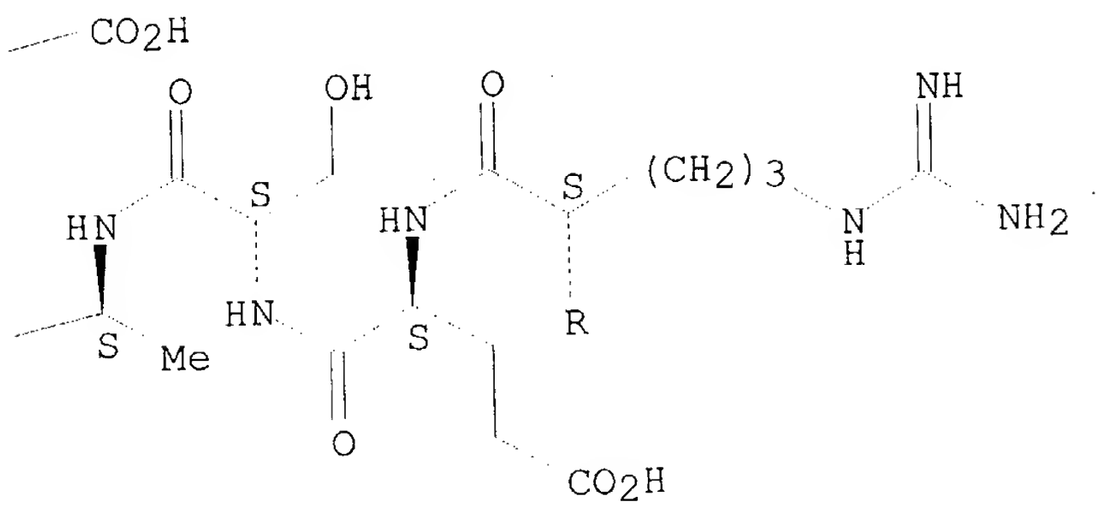
Search completed by David Schreiber x22526

Absolute stereochemistry.

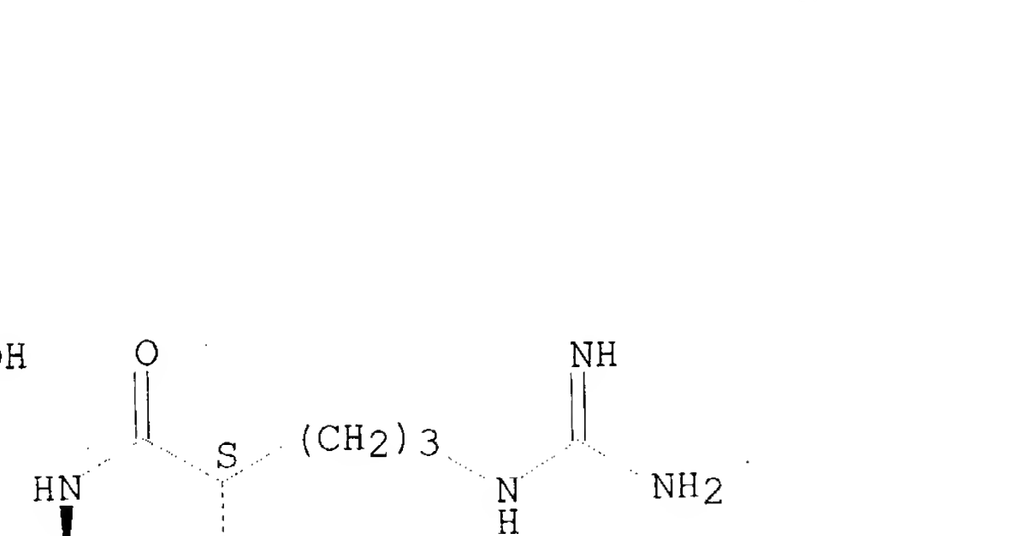
PAGE 1-A



PAGE 1-B



PAGE 2-A



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1907 TO DATE)
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

Search completed by David Schreiber x22526

```
=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 16:23:54 ON 14 JUN 2004
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)
```



```
FILE COVERS 1907 - 14 Jun 2004  VOL 140 ISS 25
FILE LAST UPDATED: 13 Jun 2004  (20040613/ED)

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.


=> d que l2
L1             1 SEA FILE=REGISTRY GGG['GLU''GLA']VR['GLU''GLA']SA['GLU''GLA']TL
                 H['GLU''GLA']IT[P'HYP']|GGG['GLU''GLA']VR['GLU''GLA']SA['GLU''G
                 LA']TLH['GLU''GLA']ITP/SQSP
L2             1 SEA FILE=HCAPLUS L1


=> d ibib abs hitrn l2 1

L2   ANSWER 1 OF 1  HCAPLUS  COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER:          2002:716012  HCAPLUS
DOCUMENT NUMBER:           137:243330
TITLE:                     Linear γ-carboxyglutamate-rich conotoxins with
                           possible therapeutic uses
INVENTOR(S):               Olivera, Baldomero M.; McIntosh, J. Michael; Garrett,
                           James E.; Walker, Craig S.; Watkins, Maren; Jones,
                           Robert M.
PATENT ASSIGNEE(S):        University of Utah Research Foundation, USA; Cognetix,
                           Inc.
SOURCE:                    PCT Int. Appl., 113 pp.
                           CODEN: PIXXD2
DOCUMENT TYPE:             Patent
LANGUAGE:                  English
FAMILY ACC. NUM. COUNT:    1
PATENT INFORMATION:

     PATENT NO.          KIND  DATE          APPLICATION NO.   DATE
     ----------------    ----  --------      ----------------  --------
     WO 2002072005        A2   20020919      WO 2002-US6863    20020307
     WO 2002072005        A3   20030123
         W:  AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
             CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
             GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
```

Search completed by David Schreiber x22526

LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH,
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ,
UA, UG, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH,
CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR,
BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

US 2003065138 A1 20030403 US 2002-92367 20020307
US 2001-273639P P 20010307
PRIORITY APPLN. INFO.:

AB The invention relates to linear γ-carboxyglutamate rich conotoxins, derivs. or pharmaceutically acceptable salts thereof, and uses thereof, including the treatment of neurol. and psychiatric disorders, such as anticonvulsant agents, as neuroprotective agents, as neuroprotective agents or for the management of pain. The invention further relates to nucleic acid sequences encoding the conopeptides and encoding propeptides, as well as the propeptides.

IT **459871-55-7**
RL: PRP (Properties)
(unclaimed sequence; linear γ-carboxyglutamate-rich conotoxins with possible therapeutic uses)

```
=> d que l7
L1             1 SEA FILE=REGISTRY GGG['GLU''GLA']VR['GLU''GLA']SA['GLU''GLA']TL
                 H['GLU''GLA']IT[P'HYP']|GGG['GLU''GLA']VR['GLU''GLA']SA['GLU''G
                 LA']TLH['GLU''GLA']ITP/SQSP
L2             1 SEA FILE=HCAPLUS L1
L3          2881 SEA FILE=HCAPLUS CONOTOXIN#
L4             3 SEA FILE=HCAPLUS L3 AND (GLUTAMICCARBOXY? OR GLUTAM?(A)CARBOXY?
                 )
L5            28 SEA FILE=HCAPLUS L3 AND (CARBOXYGLUTAM? OR CARBOXY?(A)GLUTAM?)

L6            28 SEA FILE=HCAPLUS L4 OR L5
L7            27 SEA FILE=HCAPLUS L6 NOT L2


=> d ibib abs l7 1-27

L7   ANSWER 1 OF 27  HCAPLUS  COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER:        2004:437807  HCAPLUS
TITLE:                   Determining sequences and post-translational
                         modifications of novel conotoxins in Conus
                         victoriae using cDNA sequencing and mass spectrometry
AUTHOR(S):               Jakubowski, Jennifer A.; Keays, David A.; Kelley,
                         Wayne P.; Sandall, David W.; Bingham, Jon-Paul;
                         Livett, Bruce G.; Gayler, Ken R.; Sweedler, Jonathan
                         V.
CORPORATE SOURCE:        Department of Chemistry and the Beckman Institute,
                         University of Illinois, Urbana-Champaign, IL, 61801,
                         USA
SOURCE:                  Journal of Mass Spectrometry (2004), 39(5), 548-557
                         CODEN: JMSPFJ; ISSN: 1076-5174
PUBLISHER:               John Wiley & Sons Ltd.
DOCUMENT TYPE:           Journal
LANGUAGE:                English
```

AB A combination of cDNA cloning and detailed mass spectrometric analyses was employed to identify novel **conotoxins** from Conus victoriae. Eleven **conotoxin** sequences were determined using mol. methods: one belonging to the A superfamily (Vc1.1), six belonging to the O superfamily

Search completed by David Schreiber x22526

(Vc6.1-Vc6.6) and four members of the T superfamily (Vc5.1-Vc5.4). In order to verify the sequences and identify the post-translational modifications (excluding the disulfide connectivity) of three Conus victoriae **conotoxins**, vc1a, vc5a and vc6a, deduced from sequences Vc1.1, Vc5.1, and Vc6.1, resp., liquid chromatog./electrospray ionization ion trap mass spectrometry, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and nanospray ionization ion trap mass spectrometry with collisionally induced dissociation were performed on reduced and alkylated venom fractions. We report that vc1a, the native form of α- **conotoxin** Vc1.1 (an unmodified 16 amino acid residue peptide that has notable pain-relieving capabilities), includes a hydroxyproline and a γ- **carboxyglutamate** residue. **Conotoxin** vc5a is a 10-residue peptide with two disulfide bonds and a hydroxyproline and vc6a is a 25 amino acid peptide with three disulfide bonds.

L7 ANSWER 2 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2003:952792 HCAPLUS
DOCUMENT NUMBER: 140:194777
TITLE: Efficient oxidative folding of **conotoxins** and the radiation of venomous cone snails
AUTHOR(S): Bulaj, Grzegorz; Buczek, Olga; Goodsell, Ian; Jimenez, Elsie C.; Kranski, Jessica; Nielsen, Jacob S.; Garrett, James E.; Olivera, Baldomero M.
CORPORATE SOURCE: Department of Biology, University of Utah, Salt Lake City, UT, 84112, USA
SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2003), 100(Suppl. 2), 14562-14568
CODEN: PNASA6; ISSN: 0027-8424
PUBLISHER: National Academy of Sciences
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The 500 different species of venomous cone snails (genus Conus) use small, highly structured peptides (**conotoxins**) for interacting with prey, predators, and competitors. These peptides are produced by translating mRNA from many genes belonging to only a few gene superfamilies. Each translation product is processed to yield a great diversity of different mature toxin peptides (≈50,000-100,000), most of which are 12-30 aa in length with two to three disulfide crosslinks. In vitro, forming the biol. relevant disulfide configuration is often problematic, suggesting that in vivo mechanisms for efficiently folding the diversity of **conotoxins** have been evolved by the cone snails. We demonstrate here that the correct folding of a Conus peptide is facilitated by a posttranslationally modified amino acid, γ- **carboxyglutamate**. In addition, we show that multiple isoforms of protein disulfide isomerase are major soluble proteins in Conus venom duct exts. The results provide evidence for the type of adaptations required before cone snails could systematically explore the specialized biochem. world of "microproteins" that other organisms have not been able to systematically access. Almost certainly, addnl. specialized adaptations for efficient microprotein folding are required.

REFERENCE COUNT: 29 THERE ARE 29 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 3 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2003:62579 HCAPLUS
DOCUMENT NUMBER: 139:48393
TITLE: Isolation, Structure, and Activity of GID, a Novel

Search completed by David Schreiber x22526

α4/7- **Conotoxin** with an Extended N-terminal Sequence

AUTHOR(S): Nicke, Annette; Loughnan, Marion L.; Millard, Emma L.; Alewood, Paul F.; Adams, David J.; Daly, Norelle L.; Craik, David J.; Lewis, Richard J.

CORPORATE SOURCE: Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, 4072, Australia

SOURCE: Journal of Biological Chemistry (2003), 278(5), 3137-3144
CODEN: JBCHA3; ISSN: 0021-9258

PUBLISHER: American Society for Biochemistry and Molecular Biology

DOCUMENT TYPE: Journal
LANGUAGE: English

AB Using assay-directed fractionation of Conus geographus crude venom, we isolated α- **conotoxin** GID, which acts selectively at neuronal nicotinic acetylcholine receptors (nAChRs). Unlike other neuronally selective α- **conotoxins**, α-GID has a four amino acid N-terminal tail, γ- **carboxyglutamate** (Gla), and hydroxyproline (O) residues, and lacks an amidated C terminus. GID inhibits α7 and α3β2 nAChRs with IC50 values of 5 and 3 nM, resp. and is at least 1000-fold less potent at the α1β1γδ, α3β4, and α4β4 combinations. GID also potently inhibits the α4β2 subtype (IC50 of 150 nM). Deletion of the N-terminal sequence (GIDΔ1-4) significantly decreased activity at the α4β2 nAChR but hardly affected potency at α3β2 and α7 nAChRs, despite enhancing the off-rates at these receptors. In contrast, Arg12 contributed to α4β2 and α7 activity but not to α3β2 activity. The three-dimensional structure of GID is well defined over residues 4-19 with a similar motif to other α- **conotoxins**. However, despite its influence on activity, the tail appears to be disordered in solution Comparison of GID with other α4/7- **conotoxins** which possess an NN(P/O) motif in loop II, revealed a correlation between increasing length of the aliphatic side-chain in position 10 (equivalent to 13 in GID) and greater α7 vs. α3β2 selectivity.

REFERENCE COUNT: 53 THERE ARE 53 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 4 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:12557 HCAPLUS
DOCUMENT NUMBER: 138:267664

TITLE: Expression and characterization of recombinant vitamin K-dependent γ- **glutamyl carboxylase** from an invertebrate, Conus textile

AUTHOR(S): Czerwiec, Eva; Begley, Gail S.; Bronstein, Mila; Stenflo, Johan; Taylor, Kevin; Furie, Barbara C.; Furie, Bruce

CORPORATE SOURCE: Marine Biological Laboratory, Woods Hole, MA, USA

SOURCE: European Journal of Biochemistry (2002), 269(24), 6162-6172
CODEN: EJBCAI; ISSN: 0014-2956

PUBLISHER: Blackwell Science Ltd.
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The marine snail Conus is the sole invertebrate wherein both the vitamin K-dependent carboxylase and its product, γ- **carboxyglutamic**

Search completed by David Schreiber x22526

acid, have been identified. To examine its biosynthesis of γ-**carboxyglutamic** acid, we studied the carboxylase from Conus venom ducts. The carboxylase cDNA from Conus textile has an ORF that encodes a 811-amino-acid protein which exhibits sequence similarity to the vertebrate carboxylases, with 41% identity and ≈ 60% sequence similarity to the bovine carboxylase. Expression of this cDNA in COS cells or insect cells yielded vitamin K-dependent carboxylase activity and vitamin K-dependent epoxidase activity. The recombinant carboxylase has a mol. mass of ≈ 130 kDa. The recombinant Conus carboxylase carboxylated Phe-Leu-Glu-Glu-Leu and the 28-residue peptides based on residues - 18 to + 10 of human proprothrombin and proFactor IX with Km values of 420 μM, 1.7 μM and 6 μM, resp.; the Km for vitamin K is 52 μM. The Km values for peptides based on the sequence of the **conotoxin** ε-TxIX and two precursor analogs containing 12 or 29 amino acids of the propeptide region are 565 μM, 75 μM and 74 μM, resp. The recombinant Conus carboxylase, in the absence of endogenous substrates, is stimulated up to fivefold by vertebrate propeptides but not by Conus propeptides. These results suggest two propeptide-binding sites in the carboxylase, one that binds the Conus and vertebrate propeptides and is required for substrate binding, and the other that binds only the vertebrate propeptide and is required for enzyme stimulation. The marked functional and structural similarities between the Conus carboxylase and vertebrate vitamin K-dependent γ-carboxylases argue for conservation of a vitamin K-dependent carboxylase across animal species and the importance of γ- **carboxyglutamic** acid synthesis in diverse biol. systems.

REFERENCE COUNT: 54 THERE ARE 54 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 5 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:863490 HCAPLUS
DOCUMENT NUMBER: 138:380659
TITLE: Structure of a Novel P-superfamily Spasmodic **Conotoxin** Reveals an Inhibitory Cystine Knot Motif
AUTHOR(S): Miles, Luke A.; Dy, Catherine Y.; Nielsen, Jake; Barnham, Kevin J.; Hinds, Mark G.; Olivera, Baldomero M.; Bulaj, Grzegorz; Norton, Raymond S.
CORPORATE SOURCE: NMR Laboratory, The Walter and Eliza Hall Institute of Medical Research, Parkville, 3052, Australia
SOURCE: Journal of Biological Chemistry (2002), 277(45), 43033-43040
CODEN: JBCHA3; ISSN: 0021-9258
PUBLISHER: American Society for Biochemistry and Molecular Biology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB **Conotoxin** gm9a, a putative 27-residue polypeptide encoded by Conus gloriamaris, was recently identified as a homolog of the "spasmodic peptide", tx9a, isolated from the venom of the mollusk-hunting cone shell Conus textile (M. B. Lirazan, et al. 2000). The C. gloriamaris spasmodic peptide has been synthesized, and the refolded polypeptide was shown to be biol. active using a mouse bioassay. The chemical synthesized gm9a elicited the same symptomatol. described previously for natively folded tx9a, and gm9a and tx9a were of similar potency, implying that neither the two γ- **carboxyglutamate** (Gla) residues found in tx9a (Ser8 and Ala13 in gm9a) nor Gly1 (Ser1 in gm9a) are crucial for biol. activity. We have determined the three-dimensional structure of gm9a in aqueous solution and demonstrated that the mol. adopts the well known inhibitory cystine knot

Search completed by David Schreiber x22526

motif constrained by three disulfide bonds involving Cys2-Cys16, Cys6-Cys18 and Cys12-Cys23. Based on the gm9a structure, the sites of Gla substitution in tx9a are in loops located on one surface of the mol., which is unlikely to be involved directly in receptor binding. Because this is the first structure reported for a member of the newly defined P-superfamily **conotoxins**, a comparison has been made with structurally related **conotoxins**. This shows that the structural scaffold that characterizes the P-**conotoxins** has the greatest potential for exhibiting structural diversity among the robust inhibitory cystine knot-containing **conotoxins**, a finding that has implications for functional epitope mimicry and protein engineering.

REFERENCE COUNT: 45 THERE ARE 45 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 6 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2002:777965 HCAPLUS
DOCUMENT NUMBER: 137:289027
TITLE: Alpha **conotoxin** peptides with analgesic properties
INVENTOR(S): Livett, Bruce; Khalil, Zeinab; Gayler, Kenwyn; Down, John
PATENT ASSIGNEE(S): Australia
SOURCE: PCT Int. Appl., 87 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002079236	A1	20021010	WO 2002-AU411	20020328
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
EP 1385874	A1	20040204	EP 2002-713927	20020328
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR				

PRIORITY APPLN. INFO.:
AU 2001-4094 A 20010329
WO 2002-AU411 W 20020328

OTHER SOURCE(S): MARPAT 137:289027

AB This invention relates to novel α- **conotoxin**-like peptides comprising the following sequence of amino acids: Xaa1CCSXaa2Xaa3Xaa4CXaa5Xaa6Xaa7Xaa8Xaa9Xaa10Xaa11C-NH2 in which Xaa1 is G or D; Xaa3 is proline, hydroxyproline or glutamine; each of Xaa2 to Xaa8 and Xaa11 is independently any amino acid; Xaa9 is proline, hydroxyproline or glutamine; Xaa10 is aspartate, glutamate or γ- **carboxyglutamate**; Xaa11 is optionally absent; and the C-terminus is optionally amidated, with the proviso that the peptide is not α- **conotoxin** Ep1 or α- **conotoxin** Im1. The peptides are useful in the treatment or prevention of pain, in recovery from nerve injury, and in the treatment of painful neurol. conditions such as stroke.

Search completed by David Schreiber x22526

REFERENCE COUNT: 14 THERE ARE 14 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 7 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2001:611748 HCAPLUS
DOCUMENT NUMBER: 135:190428
TITLE: Use of conantokins for treating pain
INVENTOR(S): Olivera, Baldomero M.; McIntosh, J. Michael; McCabe, R. Tyler; Layer, Richard T.; Zhou, Li-Ming
PATENT ASSIGNEE(S): University of Utah Research Foundation, USA; Cognetix, Inc.
SOURCE: U.S., 60 pp., Cont.-in-part of U.S. Ser. No. 283,277.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6277825	B1	20010821	US 1999-357141	19990720
WO 9803189	A1	19980129	WO 1997-US12652	19970721

W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.:

US 1996-684750	A2	19960722
US 1996-762377	A2	19961206
WO 1997-US12652	W	19970721
US 1999-142076	A1	19990210
US 1999-283277	A2	19990401

OTHER SOURCE(S): MARPAT 135:190428

AB The present invention is directed to the use of conantokin peptides, conantokin peptide derivs. and conantokin peptide chimeras, referred to collectively as conantokins, having 10-30 amino acids, including preferably two or more γ- **carboxyglutamic** acid residues, for the treatment of neurol. and psychiatric disorders, such as pain, e.g., as an analgesic agent.

REFERENCE COUNT: 16 THERE ARE 16 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 8 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2000:894820 HCAPLUS
DOCUMENT NUMBER: 134:349220
TITLE: Post-translational modification: A two-dimensional strategy for molecular diversity of Conus peptides
AUTHOR(S): Hooper, David; Lirazan, Marcelina B.; Schoenfeld, Robert; Cook, Brady; Cruz, Lourdes J.; Olivera, Baldomero M.; Bandyopadhyay, Pradip
CORPORATE SOURCE: Department of Biology, University of Utah, Salt Lake City, UT, 84112, USA
SOURCE: Peptides for the New Millennium, Proceedings of the American Peptide Symposium, 16th, Minneapolis, MN, United States, June 26-July 1, 1999 (2000), Meeting Date 1999, 727-729. Editor(s): Fields, Gregg B.; Tam,

Search completed by David Schreiber x22526

James P.; Barany, George. Kluwer Academic Publishers: Dordrecht, Neth.
CODEN: 69ATHX
DOCUMENT TYPE: Conference
LANGUAGE: English
AB The venomous cone snails (Conus), arguably the largest living genus of marine animals, use venom for capturing prey, defense and other purposes. The venoms contain 50-200 relatively small peptides that specifically target receptors and ion channels. A remarkable intra- and interspecific pharmacol. diversity has evolved in Conus peptides. This paper focuses on one facet of this diversity, the unprecedented variety of post-translational modifications found in these peptides.
REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 9 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2000:189695 HCAPLUS
DOCUMENT NUMBER: 132:344323
TITLE: Structure-function relationships of the NMDA receptor antagonist peptide, conantokin-R
AUTHOR(S): Blandl, T.; Warder, S. E.; Prorok, M.; Castellino, F. J.
CORPORATE SOURCE: Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN, USA
SOURCE: FEBS Letters (2000), 470(2), 139-146
CODEN: FEBLAL; ISSN: 0014-5793
PUBLISHER: Elsevier Science B.V.
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Conantokin-R (con-R) is a γ- **carboxyglutamate**-containing 27-residue neuroactive peptide present in the venom of Conus radiatus, and acts as a non-competitive antagonist of the N-methyl-D-aspartate (NMDA) receptor. This peptide features a single disulfide bond, a type of structural element found in most classes of **conotoxins**, but not in other conantokins. The NMDA receptor antagonist activity of chemical synthesized con-R was determined through an assay involving inhibition of the spermine-enhanced binding of the NMDA receptor channel blocker, [3H]MK-801, to rat brain membranes, and yielded an IC50 of 93 nM. This value represents a 2-5 times better potency than con-G or con-T, the other two characterized conantokins. CD anal. of the metal-free form of con-R is indicative of a low α-helical content. There is an increase in α-helicity upon the addition of divalent cations, such as Ca2+, Mg2+, or Zn2+. Isothermal titration calorimetry expts. showed one detectable Mg2+ binding site with a Kd of 6.5 μM, and two binding sites for Zn2+, with Kd values of 150 nM and 170 μM. Residue-specific information of the conformational state of con-R was obtained by two-dimensional 1H-NMR. Analyses of the α-proton chemical shifts, NOE patterns, and hydrogen exchange rates of the peptide indicated an α-helical conformation for residues 1-19. Synthetic con-R-derived peptide variants, containing deletions of 7 and 10 amino acid residues from the carboxy-terminus of the wild-type peptide, displayed unaltered cation binding and NMDA receptor antagonist properties. The α-helical secondary structures of the two truncation peptides were more stable than full-length con-R, as evidenced by CD measurements and reduced backbone hydrogen exchange rates. These results provide exptl. evidence that the structural elements common to the three conantokins thus far identified are the primary determinants for receptor function and cation binding/secondary structure stability.
REFERENCE COUNT: 29 THERE ARE 29 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

Search completed by David Schreiber x22526

L7 ANSWER 10 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2000:68909 HCAPLUS
DOCUMENT NUMBER: 132:330728
TITLE: The spasmodic peptide defines a new **conotoxin** superfamily
AUTHOR(S): Lirazan, Marcelina B.; Hooper, David; Corpuz, Gloria P.; Ramilo, Cecilia A.; Bandyopadhyay, Pradip; Cruz, Lourdes J.; Olivera, Baldomero M.
CORPORATE SOURCE: Department of Biology, University of Utah, Salt Lake City, UT, 84112, USA
SOURCE: Biochemistry (2000), 39(7), 1583-1588
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
AB A peptide from the venom of Conus textile that makes normal mice assume the phenotype of a well-known mutant, the spasmodic mouse, was purified and characterized. This spasmodic peptide has 27 amino acids, including two γ- **carboxyglutamate** (Gla) residues. A cDNA clone encoding the precursor for the peptide was identified; a γ-carboxylation recognition signal sequence (γ-CRS) is present in the -1 → -20 region of the peptide precursor. Both the γ-CRS and the position of the Gla residues in the mature toxin are notably different from other Gla-containing conopeptides. The spasmodic peptide has a novel disulfide framework and distinct signal sequence which together define a new P-superfamily of conopeptides. A cDNA encoding another member of the P-superfamily was identified from a different species, Conus gloriamaris.
REFERENCE COUNT: 31 THERE ARE 31 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 11 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:700553 HCAPLUS
DOCUMENT NUMBER: 132:19938
TITLE: The T-superfamily of **conotoxins**
AUTHOR(S): Walker, Craig S.; Steel, Douglas; Jacobsen, Richard B.; Lirazan, Marcelina B.; Cruz, Lourdes J.; Hooper, David; Shetty, Reshma; DelaCruz, Richard C.; Nielsen, Jacob S.; Zhou, Li Ming; Bandyopadhyay, Pradip; Craig, A. Grey; Olivera, Baldomero M.
CORPORATE SOURCE: Department of Biology, University of Utah, Salt Lake City, UT, 84112, USA
SOURCE: Journal of Biological Chemistry (1999), 274(43), 30664-30671
CODEN: JBCHA3; ISSN: 0021-9258
PUBLISHER: American Society for Biochemistry and Molecular Biology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB We report the discovery and initial characterization of the T-superfamily of **conotoxins**. Eight different T-superfamily peptides from five Conus species were identified; they share a consensus signal sequence, and a conserved arrangement of cysteine residues (- -CC- -CC-). T-superfamily peptides were found expressed in venom ducts of all major feeding types of Conus; the results suggest that the T-superfamily will be a large and diverse group of peptides, widely distributed in the 500 different Conus species. These peptides are likely to be functionally diverse; although the peptides are small (11-17 amino acids), their sequences are strikingly

Search completed by David Schreiber x22526

divergent, with different peptides of the superfamily exhibiting varying extents of post-translational modification. Of the three peptides tested for in vivo biol. activity, only one was active on mice but all three had effects on fish. The peptides that have been extensively characterized are as follows: p5a, GCCP-KQMRCCTL*; tx5a, γCCγDGW+CCTAAO; and au5a, FC-CPFIRYCCW (where γ = γ- **carboxyglutamate**, W+ = bromotryptophan, O = hydroxyproline, T = glycosylated threonine, and * = COOH-terminal amidation). We also demonstrate that the precursor of tx5a contains a functional γ-carboxylation recognition signal in the -1 to -20 propeptide region, consistent with the presence of γ-**carboxyglutamate** residues in this peptide.

REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 12 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1999:652991 HCAPLUS

DOCUMENT NUMBER: 132:74725

TITLE: Hydrophobic Amino Acids Define the Carboxylation Recognition Site in the Precursor of the γ-**Carboxyglutamic**-Acid-Containing **Conotoxin** ε-TxIX from the Marine Cone Snail Conus textile

AUTHOR(S): Bush, Kristine A.; Stenflo, Johan; Roth, David A.; Czerwiec, Eva; Harrist, Alexia; Begley, Gail S.; Furie, Barbara C.; Furie, Bruce

CORPORATE SOURCE: Center for Hemostasis and Thrombosis Research, Harvard Medical School and Beth Israel Deaconess Medical Center, Boston, MA, 02215, USA

SOURCE: Biochemistry (1999), 38(44), 14660-14666
CODEN: BICHAW; ISSN: 0006-2960

PUBLISHER: American Chemical Society

DOCUMENT TYPE: Journal

LANGUAGE: English

AB To identify the amino acid sequence of the precursor of the Gla-containing peptide, ε-TxIX, from the venom of the marine snail Conus textile, the cDNA encoding this peptide was cloned from a C. textile venom duct library. The cDNA of the precursor form of ε-TxIX encodes a 67 amino acid precursor peptide, including an N-terminal prepro-region, the mature peptide, and four residues posttranslationally cleaved from the C-terminus. To determine the role of the propeptide in γ-carboxylation, peptides were designed and synthesized based on the propeptide sequence of the Gla-containing **conotoxin** ε-TxIX and used in assays with the vitamin K-dependent γ- **glutamyl carboxylase** from C. textile venom ducts. The mature acarboxy peptide ε-TxIX was a high KM substrate for the γ-carboxylase. Synthetic peptides based on the precursor ε-TxIX were low KM substrates (5 μM) if the peptides included at least 12 residues of propeptide sequence, from -12 to -1. Leucine-19, leucine-16, asparagine-13, leucine-12, leucine-8 and leucine-4 contribute to the interaction of the pro-**conotoxin** with carboxylase since their replacement by aspartic acid increased the KM of the substrate peptide. Although the Conus propeptide and the propeptides of the mammalian vitamin K-dependent proteins show no obvious sequence homol., synthetic peptides based upon the structure of pro-ε-TxIX were intermediate KM substrates for the bovine carboxylase. The propeptide of ε-TxIX contains significant α-helix, as estimated by measurement of the CD spectra, but the region of the propeptide that plays the dominant role in directing carboxylation does not contain evidence of helical structure. These results indicate that the γ-carboxylation recognition site is defined by hydrophobic

Search completed by David Schreiber x22526

residues in the propeptide of this **conotoxin** precursor.
REFERENCE COUNT: 41 THERE ARE 41 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 13 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:404858 HCAPLUS
DOCUMENT NUMBER: 131:54035
TITLE: Gamma-conopeptide agonists for neuronal pacemaker calcium channels
INVENTOR(S): Fainzilber, Michael; Kits, Karel S.; Burlingame, Alma L.; Olivera, Baldomero M.; Walker, Craig; Walkins, Maren; Shetty, Reshma; Cruz, Lourdes J.; Imperial, Julita; Colledge, Clark
PATENT ASSIGNEE(S): University of Utah Research Foundation, USA; Vrije Universiteit; The Regents of the University of California
SOURCE: PCT Int. Appl., 59 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9930732	A1	19990624	WO 1998-US26792	19981216
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6624288	B1	20030923	US 1998-210952	19981215
CA 2314686	AA	19990624	CA 1998-2314686	19981216
AU 9920001	A1	19990705	AU 1999-20001	19981216
EP 1039923	A1	20001004	EP 1998-964743	19981216
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
JP 2002508945	T2	20020326	JP 2000-538711	19981216

PRIORITY APPLN. INFO.:
US 1997-69706P P 19971216
WO 1998-US26792 W 19981216

AB This invention relates to relatively short peptides about 25-40 residues in length, which are naturally available in minute amts. in the venom of the cone snails or analogs to the naturally available peptides, and which include three cyclizing disulfide linkages and one or more γ-**carboxyglutamate** residues. More specifically, the present invention is directed to γ-conopeptides having the general formula: Xaa1-Cys-Xaa2-Cys-Xaa3-Xaa4-Cys-Cys-Xaa5-Cys-Xaa6-Cys-Xaa7 (SEQ ID NO:1), as described herein; or having the general formula: Xaa1-Cys-Xaa2-Cys-Xaa3-Xaa4-Cys-Cys-Xaa5-Xaa6-Cys-Xaa7-Cys -Xaa8 (SEQ ID NO:2), as defined herein; or having the general formula: Xaa1-Cys-Xaa2-Cys-Xaa3-Xaa4-Xaa5-Cys-Cys-Ser-Asn-Ser-Cys-Asp -Xaa6-Cys-Xaa7 (SEQ ID NO:3), as described herein; or having the general formula: Xaa1-Cys-Xaa2-Cys-Xaa3-Xaa4-Xaa5-Cys-Cys-Ser-Asn-Ser-Cys-Asp -Xaa6-Cys-Xaa7 (SEQ ID NO:4), as described herein; or having the general formula: Xaa1-Xaa2-Cys-Xaa3-Xaa4-Phe-Xaa5-Cys-Thr-Xaa6-Ser-Xaa7-Cys-Cys-Ser-Asn-Ser-Cys-Asp-Gln-Thr-Tyr-Cys-Xaa8-Leu-Xaa9 (SEQ ID NO:5), as described herein. The

Search completed by David Schreiber x22526

invention further relates to specific γ-conopeptides, specific pro-γ-conopeptides and nucleic acids encoding the pro-γ-conopeptides. The invention also includes pharmaceutically acceptable salts of the conopeptides. These conopeptides are useful as agonists of neuronal pacemaker calcium channels.

REFERENCE COUNT: 2 THERE ARE 2 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 14 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:314480 HCAPLUS
DOCUMENT NUMBER: 131:84184
TITLE: A **conotoxin** from Conus textile with unusual posttranslational modifications reduces presynaptic Ca2+ influx
AUTHOR(S): Rigby, Alan C.; Lucas-Meunier, Estelle; Kalume, Dario E.; Czerwiec, Eva; Hambe, Bjorn; Dahlqvist, Ingrid; Fossier, Philippe; Baux, Gerard; Roepstorff, Peter; Baleja, James D.; Furie, Barbara C.; Furie, Bruce; Stenflo, Johan
CORPORATE SOURCE: Marine Biological Laboratory, Woods Hole, MA, 02543, USA
SOURCE: Proceedings of the National Academy of Sciences of the United States of America (1999), 96(10), 5758-5763
CODEN: PNASA6; ISSN: 0027-8424
PUBLISHER: National Academy of Sciences
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Cone snails are gastropod mollusks of the genus Conus that live in tropical marine habitats. They are predators that paralyze their prey by injection of venom containing a plethora of small, conformationally constrained peptides (**conotoxins**). We report the identification, characterization, and structure of a γ-**carboxyglutamic** acid-containing peptide, **conotoxin** ε-TxIX, isolated from the venom of the molluscivorous cone snail, Conus textile. The disulfide bonding pattern of the four cysteine residues, an unparalleled degree of posttranslational processing including bromination, hydroxylation, and glycosylation, define a family of **conotoxins** that may target presynaptic Ca2+ channels or act on G protein-coupled presynaptic receptors via another mechanism. This **conotoxin** selectively reduces neurotransmitter release at an Aplysia cholinergic synapse by reducing the presynaptic influx of Ca2+ in a slow and reversible fashion. The three-dimensional structure, determined by two-dimensional 1H NMR spectroscopy, identifies an electroneg. patch created by the side chains of two γ- **carboxyglutamic** acid residues that extend outward from a cavernous cleft. The glycosylated threonine and hydroxylated proline enclose a localized hydrophobic region centered on the brominated tryptophan residue within the constrained intercysteine region.

REFERENCE COUNT: 41 THERE ARE 41 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 15 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1998:87759 HCAPLUS
DOCUMENT NUMBER: 128:167715
TITLE: Preparation and anticonvulsant, neuroprotectant, and analgesic activity of conantokin peptide derivatives
INVENTOR(S): Abogadie, Fe C.; Cruz, Lourdes J.; Olivera, Baldomero M.; Walker, Craig; Colledge, Clark; Hillyard, David R.; Jimenez, Elsie; Layer, Richard T.; Zhou, Li-ming;

Search completed by David Schreiber x22526

PATENT ASSIGNEE(S): Shen, Gregory S.; et al. University of Utah Research Foundation, USA; Cognetix, Inc.; Abogadie, Fe C.; Cruz, Lourdes J.; Olivera, Baldomero M.; Walker, Craig; Colledge, Clark; Hillyard, David R.; Jimenez, Elsie; et al.

SOURCE: PCT Int. Appl., 122 pp.
CODEN: PIXXD2

DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9803541	A1	19980129	WO 1997-US12618	19970721
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
AU 9738861	A1	19980210	AU 1997-38861	19970721
AU 727196	B2	20001207		
EP 956292	A1	19991117	EP 1997-936111	19970721
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
JP 2001507924	T2	20010619	JP 1998-507104	19970721
US 6515103	B1	20030204	US 2000-142080	20000511
US 2003194729	A1	20031016	US 2003-357467	20030204

PRIORITY APPLN. INFO.:
US 1996-684742 A2 19960722
WO 1997-US12618 W 19970721
US 2000-142080 A3 20000511

OTHER SOURCE(S): MARPAT 128:167715

AB The present invention is directed to conantokin peptides, conantokin peptide derivs. and conantokin peptide chimeras (X1)m-Gly-X2-X3-X4-(X5)n-(X6)p-(X7)q [X1 = Lys-Pro-Gly-Arg-Lys, Lys-Pro-Gly-Arg-Lys-Asn; X2-X4 = independently any amino acid; X5 = peptide containing 1-7 amino acid residues; X6 = peptide containing 1-4 amino acid residues; X7 = peptide containing 1-12 amino acid residues; m, n, p, q = independently 0, 1, with the proviso that if m = 1 then n = p = q = 0], referred to collectively as conantokins, having 10-30 amino acids, including preferably two or more γ- **carboxyglutamic** acid (Gla) residues. The conantokins are useful for the treatment of neurol. and psychiatric disorders, such as anticonvulsant agents, neuroprotective agents or analgesic agents. The sequence of sleeper-I peptide isolated from conus radiatus was identified as H-Gly-Glu-Gla-Gla-Val-Ala-Lys-Met-Ala-Ala-Gla-Leu-Ala-Arg-Gla-Asn-Ile-Ala-Lys-Gly-Cys-Lys-Val-Asn-Cys-Tyr-Pro-OH (Cys-Cys)-cyclic disulfide and designated as conantokin R (for radiatus). Isolation of DNA encoding conantokins is also described. A variety of conantokin R derivs. and chimeras were prepared and tested for NMDA inhibitory activity using a spermine-stimulated [3H]MK-801 binding assay in female rats. Other conantokins, including conantokin R, were tested for anticonvulsant and antiparkinsonian activities, as well as biol. stability.

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 16 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN

Search completed by David Schreiber x22526

ACCESSION NUMBER: 1998:56257 HCAPLUS
DOCUMENT NUMBER: 128:111763
TITLE: γ- **Conotoxin**-PnVIIA, A γ-**Carboxyglutamate**-Containing Peptide Agonist of Neuronal Pacemaker Cation Currents
AUTHOR(S): Fainzilber, Michael; Nakamura, Takemichi; Lodder, Johannes C.; Zlotkin, Eliahu; Kits, Karel S.; Burlingame, Alma L.
CORPORATE SOURCE: Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, 76100, Israel
SOURCE: Biochemistry (1998), 37(6), 1470-1477
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
AB A novel γ- **carboxyglutamate**-containing peptide, designated γ- **conotoxin**-PnVIIA, is described from the venom of the molluscivorous snail Conus pennaceus. γPnVIIA triggers depolarization and firing of action potential bursts in the caudodorsal neurons of Lymnaea. This effect is due to activation or enhancement of a slow inward cation current that may underly endogenous bursting activity of these neurons. The amino acid sequence of γPnVIIA was determined as DCTSWFGRCTVNSγCCSNSCDQTYCγLYAFOS (where γ is γ-**carboxyglutamate**, O is trans-4-hydroxyproline), thus γPnVIIA belongs to the six cysteine four loop structural family of **conotoxins**, and is most homologous to the previously described excitatory **conotoxin**-TxVIIA. Interestingly, TxVIIA did not induce action potentials in Lymnaea caudodorsal neurons. γPnVIIA is the prototype of a new class of γ- **conotoxins** that will provide tools for the study of voltage-gated pacemaker channels, which underly bursting processes in excitable systems.
REFERENCE COUNT: 37 THERE ARE 37 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 17 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1998:6405 HCAPLUS
DOCUMENT NUMBER: 128:19564
TITLE: Role of γ- **Carboxyglutamic** Acid in the Calcium-Induced Structural Transition of Conantokin G, a **Conotoxin** from the Marine Snail Conus geographus
AUTHOR(S): Rigby, Alan C.; Baleja, James D.; Li, Leping; Pedersen, Lee G.; Furie, Barbara C.; Furie, Bruce
CORPORATE SOURCE: Marine Biological Laboratory, Woods Hole, MA, 02543, USA
SOURCE: Biochemistry (1997), 36(50), 15677-15684
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
AB To investigate the role of γ- **carboxyglutamic** acid (Gla)in the calcium-induced structural transition of conantokin G, we determined the three-dimensional structure of the conantokin G/Ca2+ complex by two-dimensional 1H NMR spectroscopy and compared it to the high-resolution structure of conantokin G in the absence of metal ions [Rigby et al. (1997) Biochem. 36, 6906]. Complete resonance assignments were made by two dimensional 1H NMR spectroscopy at pH 5.6 in the presence of saturating amts. of Ca2+. Distance geometry and simulated annealing methods were used to derive 23 convergent structures from a set of 302 interproton

Search completed by David Schreiber x22526

distance restraints and two torsion angle measurements. A high-resolution structure, with the backbone root mean square deviation to the geometric average of the 23 structures of 0.6±0.1 Å, contains a linear α-helix from Gla 3 to Lys 15. Gla residues 3, 7, 10, and 14 are aligned in a linear array on one face of the helix. A genetic algorithm was applied to determine the calcium positions in conantokin G, and the conantokin G/Ca2+ complex refined by mol. simulation. Upon binding of Ca2+ to γ- **carboxyglutamic** acid, conantokin G undergoes a conformational transition from a distorted curvilinear 310 helix to a linear α-helix. Occupancy of the metal binding sites, defined by γ- **carboxyglutamic** acids, results in formation of a calcium-carboxylate network that linearizes the helix and exposes the hydrophobic amino acids on the opposite face of the helix.

REFERENCE COUNT: 37 THERE ARE 37 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 18 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1997:614601 HCAPLUS
DOCUMENT NUMBER: 127:258791
TITLE: Three-Dimensional Structure of a γ-**Carboxyglutamic** Acid-Containing **Conotoxin**, Conantokin G, from the Marine Snail Conus geographus: The Metal-Free Conformer. [Erratum to document cited in CA126:302496]
AUTHOR(S): Rigby, Alan C.; Baleja, James D.; Furie, Barbara C.; Furie, Bruce
CORPORATE SOURCE: Marine Biological Laboratory, Woods Hole, MA, 02543, USA
SOURCE: Biochemistry (1997), 36(40), 12394
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Under Discussion, the comment that differences in the secondary structure in independent studies of the apoconantokin G might be due to the use of a form of conantokin G that lacks the C-terminal amide is incorrect since the peptides by Prorok et al. [Prorok, M., Warder, S. E., Blandl, T., and Castellino, F. J. (1996) Biochem. 35, 16528-16534] is amidated at the C-terminus.

L7 ANSWER 19 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1997:329336 HCAPLUS
DOCUMENT NUMBER: 126:302496
TITLE: Three-Dimensional Structure of a γ-**Carboxyglutamic** Acid-Containing **Conotoxin**, Conantokin G, from the Marine Snail Conus geographus: The Metal-Free Conformer
AUTHOR(S): Rigby, Alan C.; Baleja, James D.; Furie, Barbara C.; Furie, Bruce
CORPORATE SOURCE: Marine Biological Laboratory, Woods Hole, MA, 02543, USA
SOURCE: Biochemistry (1997), 36(23), 6906-6914
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
AB To gain insight into the role of γ- **carboxyglutamic** acid (Gla) in the structure of the title peptide, we determined the three-dimensional structure of conantokin G by 1H NMR and compared its

Search completed by David Schreiber x22526

structure to other **conotoxins** and to the γ-**carboxyglutamic** acid-containing regions of the vitamin K-dependent blood-clotting proteins. Complete resonance assignments were made by two-dimensional 1H NMR spectroscopy in the absence of metal ions. NOE cross-peaks dαN, dNN, and dβN provided interproton distance information, and vicinal spin-spin coupling consts. 3JHNα were used to calculate Φ torsion angles. Distance geometry and simulated annealing methods were used to derive 20 convergent structures from a set of 227 interproton distance restraints and 13 torsion angle measurements. The backbone rmsd to the geometric average for 20 final structures is 0.8±0.1 Å. Conantokin G consists of a structured region commencing at Gla 3 and extending through arginine 13. This structure includes a partial loop centered around Gla 3 and Gla 4, a distorted type I turn between glutamine 6 and glutamine 9, and two type I turns involving Gla 10, leucine 11, and isoleucine 12 and arginine 13. Together, these two turns define approx. 1.6 turns of a distorted 310 helix. This is the first structure determined of a γ- **carboxyglutamic** acid-containing polypeptide that is not a member of the blood-clotting family of proteins. The observed structure possesses structural elements similar to those seen in the disulfide-linked **conotoxins**.

L7 ANSWER 20 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1996:158001 HCAPLUS
DOCUMENT NUMBER: 124:223305
TITLE: Mass spectrometric-based revision of the structure of a cysteine-rich peptide toxin with γ-**carboxyglutamic** acid, TxVIIA, from the sea snail, Conus textile
AUTHOR(S): Nakamura, Takemichi; Yu, Zhonghua; Fainziler, Michael; Burlingame, Alma L.
CORPORATE SOURCE: Dep. Pharmaceutical Chemistry, Univ. California, San Francisco, CA, 94143-0446, USA
SOURCE: Protein Science (1996), 5(3), 524-30
CODEN: PRCIEI; ISSN: 0961-8368
PUBLISHER: Cambridge University Press
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The authors have recently reinvestigated the title toxin employing some of the most novel techniques in mass spectrometry. The authors now report a revised structure based primarily on high-energy collision-induced dissociation anal. of the two Asp17-N peptides of the reduced, pyridinylethyl derivative representing the entire sequence using matrix-assisted laser desorption ionization (MALDI) as CGGYSTYCγVDSγCCSDNCVRSCTLF-NH2 (γ,γ- **carboxyglutamic** acid or Gla). The N-terminus of the previous sequence was incorrect, apparently due to a side reaction of reduction and alkylation, which led to the erroneous assignment of Trp for the N-terminal residue. In addition, the last two C-terminal amino acids and the C-terminal amidation had not been detected. Also, a combination of electrospray ionization mass spectrometry and pos. and neg. ion MALDI mass spectrometry provided information on the mol. wts. of the native and derivatized toxin and presence of two Gla residues. Thus, TxVIIA does not have an "unusual" sequence as previously reported, but in fact belongs to the conserved Cys framework for ω- and δ- **conotoxins**. However, the four net neg. charges with the cysteine-rich structure of this revised sequence is highly unusual for conopeptides.

L7 ANSWER 21 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1995:776763 HCAPLUS

Search completed by David Schreiber x22526

DOCUMENT NUMBER: 124:30339
TITLE: Synthesis and disulfide structure determination of **conotoxin** GS, a γ-**carboxyglutamic** acid-containing neurotoxic peptide
AUTHOR(S): Nakao, Masayuki; Nishiuchi, Yuji; Nakata, Makoto; Watanabe, Takushi X.; Kimura, Terutoshi; Sakakibara, Shumpei
CORPORATE SOURCE: Peptide Inst., Inc., Osaka, 562, Japan
SOURCE: Letters in Peptide Science (1995), 2(1), 17-26
CODEN: LPSCEM; ISSN: 0929-5666
PUBLISHER: ESCOM
DOCUMENT TYPE: Journal
LANGUAGE: English

AB **Conotoxin** GS, a γ- **carboxyglutamic** acid (Gla)-containing neurotoxic peptide composed of 34 amino acid residues with one Gla residue and three intramol. disulfide bonds, was synthesized in solution by the Boc strategy, using the cyclohexyl group to protect the γ,γ-dicarboxyl functional side chain of the Gla residue. All of the protecting groups were removed by the HF procedure. During the synthesis, the Gla residue was completely stable and no decarboxylated product was observed The free peptide was subjected to the oxidative folding reaction. The reaction proceeded almost quant. in the presence of reduced and oxidized glutathione; however, no product was formed in the absence of redox reagents concomitant with the formation of disulfide isomers or intermediates. The final product was confirmed to be identical to natural **conotoxin** GS on reversed phase- and ion exchange-HPLC as well as capillary zone electrophoresis. The disulfide structure of synthetic **conotoxin** GS was determined by gas-phase sequencing and mass spectrometry of its proteolytic fragments and was found to be identical to those of other ω- **conotoxins**. The major disulfide isomer obtained during the oxidative folding reaction without redox reagents was determined in the same manner. To clarify the role of the Gla residue and the disulfide structure in the **conotoxin** GS mol., decarboxylated **conotoxin** GS and its disulfide isomer were also synthesized, and the neurotoxic activities and CD spectra of these peptides were compared with those of **conotoxin** GS and its disulfide isomer. The results showed that the correct disulfide structure was necessary for expression of the toxicity; however, the presence of the Gla residue was not a prerequisite for both the activity and the calcium-dependent conformational transition.

L7 ANSWER 22 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1993:619430 HCAPLUS
DOCUMENT NUMBER: 119:219430
TITLE: Polyamine-like actions of peptides derived from conantokin-G, an N-methyl-D-aspartate (NMDA) antagonist
AUTHOR(S): Chandler, Paulette; Pennington, Michael; Maccecchini, Maria Luisa; Nashed, Nashaat T.; Skolnick, Phil
CORPORATE SOURCE: Lab. Neurosci., Natl. Inst. Diabetes Dig. and Kidney Dis., Bethesda, MD, 20892, USA
SOURCE: Journal of Biological Chemistry (1993), 268(23), 17173-8
CODEN: JBCHA3; ISSN: 0021-9258
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Substitution of the highly conserved γ- **carboxyglutamate** residues as well as modification of the N and C termini of conantokin-G

Search completed by David Schreiber x22526

abolished the inhibition of polyamine responses at the NMDA receptor complex. However, several of these modified polypeptides closely mimicked the neurochem. profile of polyamines at the NMDA receptor complex. One of these derivs., Tyr0-conantokin-G, was found to be the most potent compound exhibiting polyamine-like actions at the NMDA receptor complex described to date, .apprx.7-fold more potent than spermine. CD studies demonstrate a significant α-helical content in conantokin-G (27% in aqueous medium). However, this α-helicity is not sufficient for the NMDA antagonist action of the parent peptide and is neither necessary nor sufficient for the polyamine-like behavior of several conantokin-G analogs. The modified conantokin-G derivs. described in this report should be useful probes for examining the role of both polyamines and the polyamine recognition site in the operation of the NMDA receptor complex.

L7 ANSWER 23 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1992:53280 HCAPLUS
DOCUMENT NUMBER: 116:53280
TITLE: Mollusk-specific toxins from the venom of Conus textile neovicarius
AUTHOR(S): Fainzilber, Michael; Gordon, Dalia; Hasson, Arik; Spira, Micha E.; Zlotkin, Eliahu
CORPORATE SOURCE: Dep. Zool., Hebrew Univ., Jerusalem, 91904, Israel
SOURCE: European Journal of Biochemistry (1991), 202(2), 589-95
CODEN: EJBCAI; ISSN: 0014-2956
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Three peptide toxins exhibiting strong paralytic activity to mollusks, but with no paralytic effects on arthropods or vertebrates, were purified from the venom of the molluscivorous snail C. textile neovicarius from the Red Sea. The amino acid sequences of these mollusks specific toxins are: TxIA, WCKQSGEMCNLLDQNCCDGYCIVLVCT (identical to the so called King Kong peptide); TxIB, WCKQSGEMCNVLDQNCCDGYCIVFVCT; TxIIA, WGGYSTYCγVDSγCCSDNCVRSYCT (γ = γ-**carboxyglutamate**). There is a similarity of the Cys framework of these toxins to that of the ω- **conotoxins**; however, their net neg. charges, high content of hydrophobic residues, and uneven number of Cys residues in TxIIA are highly unusual for **conotoxins**. When assayed on isolated cultured Aplysia neurons, all 3 toxins induced membrane depolarization and spontaneous repetitive firing. The TxI toxins also induce a marked prolongation of the action potential duration, which is Na-dependent. These effects differ significantly from the blocking activities of piscivorous venom **conotoxins**. These mollusk-specific **conotoxins** may, therefore, serve as new and selective probes for ion-channel functions in molluskan neuronal systems.

L7 ANSWER 24 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:402314 HCAPLUS
DOCUMENT NUMBER: 113:2314
TITLE: Conantokin-T. A γ- **carboxyglutamate** containing peptide with N-methyl-D-aspartate antagonist activity
AUTHOR(S): Haack, Julie A.; Rivier, Jean; Parks, Thomas N.; Mena, E. Edward; Cruz, Lourdes J.; Olivera, Baldomero M.
CORPORATE SOURCE: Dep. Biol., Univ. Utah, Salt Lake City, UT, 84112, USA
SOURCE: Journal of Biological Chemistry (1990), 265(11), 6025-9
CODEN: JBCHA3; ISSN: 0021-9258
DOCUMENT TYPE: Journal

Search completed by David Schreiber x22526

LANGUAGE: English
AB Conantokin-T, a 21-amino acid peptide which induces sleep-like symptoms in young mice, was purified from the venom of the fish-hunting cone snail, Conus tulipa. The amino acid sequence of the peptide was determined and verified by chemical synthesis. The peptide has 4 residues of the modified amino acid, γ- **carboxyglutamate** (Gla). The sequence of the peptide is: Gly-Glu-Gla-Gla-Tyr-Gln-Lys-Met-Leu-Gla-Asn-Leu-Arg-Gla-Ala-Glu-Val-Lys-Lys-Asn-Ala-NH2. Conantokin-T inhibits N-methyl-D-aspartate (NMDA) receptor-mediated Ca influx in central nervous system neurons. Like conantokin-G (a homologous Conus peptide with recently identified NMDA antagonist activity), conantokin-T has NMDA antagonist activity. A sequence comparison of conantokins-T and -G identifies the 4 Gla residues and the N-terminal dipeptide sequence as potential key elements for the biol. activity of this peptide.

L7 ANSWER 25 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1988:487806 HCAPLUS
DOCUMENT NUMBER: 109:87806
TITLE: A novel sodium channel inhibitor from Conus geographus: purification, structure, and pharmacological properties
AUTHOR(S): Yanagawa, Yuchio; Abe, Teruo; Satake, Mei; Odani, Shoji; Suzuki, Junichi; Ishikawa, Kiichi
CORPORATE SOURCE: Sch. Med., Niigata Univ., Niigata, Japan
SOURCE: Biochemistry (1988), 27(17), 6256-62
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
AB A novel toxin, tentatively named **conotoxin** GS (CGS), was isolated from a marine snail, C. geographus. CGS existed as a single polypeptide chain, consisting of 34 amino acid residues, crosslinked by 3 disulfide bonds. Its amino acid sequence was Ala-Cys-Ser-Gly-Arg-Gly-Ser-Arg-Cys-Hyp-Hyp-Gln-Cys-Cys-Met-Gly-Leu-Arg-Cys-Gly-Arg-Gly-Asn-Pro-Gln-Lys-Cys-Ile-Gly-Ala-His-Gla-Asp-Val (Gla = 4-**carboxyglutamic** acid). In competition expts., CGS inhibited the bindings of [3H]Lys-tetrodotoxin ([3H]Lys-TTX) and [3H]propionylconotoxin GIIIA to Electrophorus electricus electroplax membranes, with Ki values of 34 and 24 nM, resp. CGS inhibited the binding of [3H]Lys-TTX (1 nM) to rat skeletal muscle homogenates with a median inhibitory concentration value of 880 nM, but showed very little effect on this binding to the rat brain P2 fraction at 10 μM. Thus, CGS belongs to the same group of Na channel inhibitors as TTX, saxitoxin, and μ- **conotoxins**. Although CGS, like the μ- **conotoxins**, is a pharmacol. probe for distinguishing between neuronal and muscle Na channel subtypes, the homol. in the sequences of CGS and μ- **conotoxins** is very limited.

L7 ANSWER 26 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1988:449968 HCAPLUS
DOCUMENT NUMBER: 109:49968
TITLE: A model for "sleeper peptide" (**conotoxin** GV) and other CLA-containing molecules
AUTHOR(S): Gray, W. R.; Olivera, B. M.; Cruz, L. J.; Rivier, J.
CORPORATE SOURCE: Biol. Dep., Univ. Utah, Salt Lake City, UT, USA
SOURCE: Peptide Chemistry (1988), Volume Date 1987 105-13
CODEN: PECHDP; ISSN: 0388-3698
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Conus Snail sleeper peptide (**conotoxin** GV), osteocalcin, and blood-clotting proteins share some structural features in their γ-

Search completed by David Schreiber x22526

carboxyglutamate (Gla)-containing segments. Gla residues are embedded in regions that intrinsically tend toward α-helical and are restricted to 1 side of the helix. It is proposed that Ca2+ binding occurs preferentially between 2 Gla residues on adjacent turns of the helix, stabilizing it against electrostatic disruption. The α-helical form is suggested as the biol. active conformation, whether the peptide acts as a monomer or dimer, or whether it acts directly as a receptor or via binding to phospholipid membrane surfaces.

L7 ANSWER 27 OF 27 HCAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1988:1901 HCAPLUS
DOCUMENT NUMBER: 108:1901
TITLE: Total synthesis and further characterization of the γ- **carboxyglutamate**-containing "sleeper" peptide from Conus geographus venom
AUTHOR(S): Rivier, Jean; Galyean, Robert; Simon, Lajos; Cruz, Lourdes J.; Olivera, Baldomero M.; Gray, William R.
CORPORATE SOURCE: Clayton Found. Lab. Pept. Biol., Salk Inst., La Jolla, CA, 92037, USA
SOURCE: Biochemistry (1987), 26(26), 8508-12
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The total synthesis of the γ- **carboxyglutamic** acid (Gla)-containing sleeper peptide (Gly-Glu-Gla-Gla-Leu-Gln-Gla-Asn-Gln-Gla-Leu-Ile-Arg-Gla-Lys-Ser-Asn-NH2) from C. geographus is described. A new strategy for the synthesis of the acid-sensitive peptide amides was developed, which allowed complete deprotection and cleavage of the L-Gla-containing peptide from the 2,4-dimethoxybenzhydrylamine resin. Synthetic sleeper peptide, after preparative HPLC purification, was identical with the native peptide by all criteria (coelution expts. on HPLC, sequence anal., and biol. activity). In addition, a developmental switch in the behavioral symptoms induced by the peptide after intracerebral administration in mice was documented. At low doses of the peptide (4-30 pmol/g), a sleeplike state was induced in mice under 2 wk old; in contrast, older mice became markedly hyperactive. It is proposed that, in the presence of Ca2+, the sleeper peptide assumes an α-helical configuration in which all the Gla-residues are located on the same side of the α-helix.

Search completed by David Schreiber x22526

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:12:24 ; Search time 5.53488 Seconds
(without alignments)
158.565 Million cell updates/sec

Title: US-10-092-367-138
Perfect score: 72
Sequence: 1 GGGXVRXSAXTLHXITP 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	66.7	1441	4	US-09-252-991A-28143	Sequence 28143, A
2	45	62.5	415	4	US-09-252-991A-31348	Sequence 31348, A
3	44	61.1	206	2	US-08-531-525-50	Sequence 50, Appl
4	44	61.1	206	2	US-08-718-270A-50	Sequence 50, Appl
5	44	61.1	308	4	US-09-328-352-4235	Sequence 4235, Ap
6	43	59.7	187	3	US-09-078-317-11	Sequence 11, Appl
7	43	59.7	188	2	US-08-531-525-47	Sequence 47, Appl
8	43	59.7	188	2	US-08-718-270A-47	Sequence 47, Appl
9	43	59.7	204	3	US-09-078-317-14	Sequence 14, Appl
10	43	59.7	204	4	US-09-454-818-14	Sequence 14, Appl
11	43	59.7	210	4	US-09-053-374A-7	Sequence 7, Appli
12	43	59.7	213	4	US-09-503-505A-3	Sequence 3, Appli
13	43	59.7	215	2	US-08-531-525-49	Sequence 49, Appl
14	43	59.7	215	2	US-08-718-270A-49	Sequence 49, Appl
15	43	59.7	791	1	US-08-394-880B-2	Sequence 2, Appli
16	42	58.3	349	4	US-09-252-991A-24644	Sequence 24644, A
17	42	58.3	350	4	US-09-821-736-2	Sequence 2, Appli
18	42	58.3	538	4	US-09-489-039A-7647	Sequence 7647, Ap
19	42	58.3	910	4	US-09-252-991A-24095	Sequence 24095, A
20	41	56.9	136	4	US-09-252-991A-20192	Sequence 20192, A
21	41	56.9	167	4	US-09-904-615-125	Sequence 125, App
22	41	56.9	241	4	US-09-489-039A-7795	Sequence 7795, Ap
23	41	56.9	402	4	US-09-543-681A-7141	Sequence 7141, Ap
24	41	56.9	438	4	US-09-252-991A-27582	Sequence 27582, A
25	41	56.9	455	4	US-09-252-991A-16635	Sequence 16635, A
26	41	56.9	1706	4	US-09-252-991A-31760	Sequence 31760, A
27	40.5	56.2	1005	1	US-08-089-986-3	Sequence 3, Appli
28	40.5	56.2	1005	1	US-08-478-585-3	Sequence 3, Appli
29	40.5	56.2	1005	1	US-08-717-312-3	Sequence 3, Appli
30	40.5	56.2	1005	2	US-08-266-408-3	Sequence 3, Appli
31	40.5	56.2	1005	5	PCT-US94-07886-3	Sequence 4099, Ap
32	40	55.6	142	4	US-09-621-976-4099	Sequence 38, Appl
33	40	55.6	179	2	US-08-531-525-38	Sequence 38, Appl
34	40	55.6	179	2	US-08-718-270A-38	Sequence 39, Appl
35	40	55.6	183	2	US-08-531-525-39	Sequence 39, Appl
36	40	55.6	183	2	US-08-718-270A-39	Sequence 8, Appli
37	40	55.6	183	4	US-09-482-520A-8	Sequence 117, App
38	40	55.6	210	4	US-09-247-155-117	Sequence 50, Appl
39	40	55.6	306	1	US-08-233-788A-50	Sequence 4536, Ap
40	40	55.6	328	4	US-09-328-352-4536	Sequence 3957, Ap
41	40	55.6	415	4	US-09-134-001C-3957	Sequence 19516, A
42	40	55.6	453	4	US-09-252-991A-19516	Sequence 29141, A
43	40	55.6	607	4	US-09-252-991A-29141	Sequence 9, Appli
44	40	55.6	669	3	US-08-941-445A-9	Sequence 25918, A
45	40	55.6	757	4	US-09-252-991A-25918	

ALIGNMENTS

RESULT 1
US-09-252-991A-28143
; Sequence 28143, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28143
; LENGTH: 1441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28143

Query Match 66.7%; Score 48; DB 4; Length 1441;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              | |:|::||:|||:|
Db       1360 GEGDVQGSAATLHTI 1374
```

RESULT 2
US-09-252-991A-31348
; Sequence 31348, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31348
; LENGTH: 415
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31348

Query Match 62.5%; Score 45; DB 4; Length 415;
Best Local Similarity 53.3%; Pred. No. 42;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
Db 189 GAGPVRASAVLLHPM 203

RESULT 3
US-08-531-525-50
; Sequence 50, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Geodia cydonium
US-08-531-525-50

Query Match 61.1%; Score 44; DB 2; Length 206;
Best Local Similarity 53.3%; Pred. No. 26;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
Db 9 GGGLVGKSALTLQLV 23

RESULT 4
US-08-718-270A-50
, Sequence 50, Application US/08718270A
, Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Geodia cydonium
US-08-718-270A-50

Query Match 61.1%; Score 44; DB 2; Length 206;
Best Local Similarity 53.3%; Pred. No. 26;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
Db 9 GGGLVGKSALTLQLV 23

RESULT 5
US-09-328-352-4235
; Sequence 4235, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4235
; LENGTH: 308

```
   TYPE: PRT
   ORGANISM: Acinetobacter baumannii
US-09-328-352-4235

  Query Match              61.1%;  Score 44;  DB 4;  Length 308;
  Best Local Similarity    37.5%;  Pred. No. 42;
  Matches    6;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :: : ||::|
Db        240 GGGIINHTIPLLHHVT 255


RESULT 6
US-09-078-317-11
  Sequence 11, Application US/09078317
  Patent No. 6017710
   GENERAL INFORMATION:
     APPLICANT:  Allen, Maxine J.
     APPLICANT:  Rutter, Marc
     APPLICANT:  Buckler, Alan J.
     TITLE OF INVENTION:  RAQ Genes and Their Uses
     NUMBER OF SEQUENCES:  16
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Bozicevic & Reed, LLP
       STREET:  285 Hamilton Ave, Suite 200
       CITY:  Palo Alto
       STATE:  CA
       COUNTRY:  USA
       ZIP:  94301
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Diskette
       COMPUTER:  IBM Compatible
       OPERATING SYSTEM:  DOS
       SOFTWARE:  FastSEQ for Windows Version 2.0
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/09/078,317
       FILING DATE:  13-MAY-1998
       CLASSIFICATION:
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:
       FILING DATE:
     ATTORNEY/AGENT INFORMATION:
       NAME:  Francis, Carol L
       REGISTRATION NUMBER:  36,513
       REFERENCE/DOCKET NUMBER:  SEQ-18P
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  650-327-3400
       TELEFAX:  650-327-3231
       TELEX:
    INFORMATION FOR SEQ ID NO:  11:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  187 amino acids
       TYPE:  amino acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
     MOLECULE TYPE:  No. 6017710e
US-09-078-317-11

  Query Match              59.7%;  Score 43;  DB 3;  Length 187;
  Best Local Similarity    53.3%;  Pred. No. 33;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 7
US-08-531-525-47
  Sequence 47, Application US/08531525
  Patent No. 5840683
   GENERAL INFORMATION:
     APPLICANT:  Hlavka, Joseph J.
     APPLICANT:  Pincus, Matthew R.
     APPLICANT:  No. 5840683le, John F.
     APPLICANT:  Abajian, Henry B.
     APPLICANT:  Kende, Andrew S.
     TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
     TITLE OF INVENTION:  of P21 Ras
     NUMBER OF SEQUENCES:  52
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Greenlee and Winner, P.C.
       STREET:  5370 Manhattan Circle, Suite 201
       CITY:  Boulder
       STATE:  Colorado
       COUNTRY:  US
       ZIP:  80303
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/531,525
       FILING DATE:  21-SEP-1995
       CLASSIFICATION:  530
     ATTORNEY/AGENT INFORMATION:
       NAME:  Ferber, Donna M.
       REGISTRATION NUMBER:  33,878
       REFERENCE/DOCKET NUMBER:  37-94
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (303) 499-8080
       TELEFAX:  (303) 499-8089
    INFORMATION FOR SEQ ID NO:  47:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  188 amino acids
       TYPE:  amino acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
     MOLECULE TYPE:  protein
     HYPOTHETICAL:  NO
     ORIGINAL SOURCE:
       ORGANISM:  Dictyostelium discoideum
US-08-531-525-47

  Query Match              59.7%;  Score 43;  DB 2;  Length 188;
  Best Local Similarity    53.3%;  Pred. No. 33;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db          9 GGGGVGKSALTIQLI 23


RESULT 8
US-08-718-270A-47
; Sequence 47, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/718,270A
;       FILING DATE:  20-SEP-1996
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/531,525
;       FILING DATE:  21-SEP-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/004,091
;       FILING DATE:  21-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ferber, Donna M.
;       REGISTRATION NUMBER:  33,878
;       REFERENCE/DOCKET NUMBER:  78-95
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303) 499-8080
;       TELEFAX:  (303) 499-8089
;   INFORMATION FOR SEQ ID NO:  47:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  188 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Dictyostelium discoideum
JS-08-718-270A-47

  Query Match             59.7%;  Score 43;  DB 2;  Length 188;
  Best Local Similarity   53.3%;  Pred. No. 33;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db          9 GGGGVGKSALTIQLI 23


RESULT 9
US-09-078-317-14
; Sequence 14, Application US/09078317
; Patent No. 6017710
;  GENERAL INFORMATION:
;    APPLICANT:  Allen, Maxine J.
;    APPLICANT:  Rutter, Marc
;    APPLICANT:  Buckler, Alan J.
;    TITLE OF INVENTION:  RAQ Genes and Their Uses
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Ave, Suite 200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/078,317
;      FILING DATE:  13-MAY-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
,      APPLICATION NUMBER:
,      FILING DATE:
,    ATTORNEY/AGENT INFORMATION:
,      NAME:  Francis, Carol L
;      REGISTRATION NUMBER:  36,513
;      REFERENCE/DOCKET NUMBER:  SEQ-18P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-327-3400
;      TELEFAX:  650-327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6017710e
US-09-078-317-14

  Query Match             59.7%;  Score 43;  DB 3;  Length 204;
  Best Local Similarity   53.3%;  Pred. No. 37;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         21 GGGGVGKSALTIQFI 35


RESULT 10
US-09-454-818-14
; Sequence 14, Application US/09454818
; Patent No. 6383792
; GENERAL INFORMATION:
;  APPLICANT:  Allen, Maxine J.
;  APPLICANT:  Rutter, Marc
;  APPLICANT:  Buckler, Alan J.
;  TITLE OF INVENTION:  RAQ Genes and Their Uses
;  FILE REFERENCE:  AXYS-018DIV
;  CURRENT APPLICATION NUMBER: US/09/454,818
;  CURRENT FILING DATE:  1999-12-03
;  PRIOR APPLICATION NUMBER:  09/078,317
;  PRIOR FILING DATE:  1998-05-13
;  NUMBER OF SEQ ID NOS:  16
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-454-818-14

  Query Match             59.7%;  Score 43;  DB 4;  Length 204;
  Best Local Similarity   53.3%;  Pred. No. 37;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         21 GGGGVGKSALTIQFI 35


RESULT 11
US-09-053-374A-7
; Sequence 7, Application US/09053374A
; Patent No. 6462177
;  GENERAL INFORMATION:
;    APPLICANT:  YEN, KWANG-MU
;    TITLE OF INVENTION:  MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/053,374A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  COOK, ROBERT R.
;       REGISTRATION NUMBER:  31,602
;       REFERENCE/DOCKET NUMBER:  A-514
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  210 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
JS-09-053-374A-7

  Query Match             59.7%;  Score 43;  DB 4;  Length 210;
  Best Local Similarity   53.3%;  Pred. No. 38;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         28 GGGGVGKSALTIQFI 42


RESULT 12
JS-09-503-505A-3
; Sequence 3, Application US/09503505A
; Patent No. 6387688
; GENERAL INFORMATION:
;  APPLICANT: SHISHIDO, KAZUO
;  APPLICANT:  KAJIWARA, SUSUMU
;  APPLICANT:  TSUKAMOTO AKIRA
;  TITLE OF INVENTION: DNA FRAGMENTS HAVING BASIDIOMYCETE-DERIVED PROMOTER
;  TITLE OF INVENTION:  ACTIVITY AND EXPRESSION OF FOREIGN GENES UNDER
;  TITLE OF INVENTION:  CONTROL OF THE PROMOTER ACTIVITY
;  FILE REFERENCE: 04853.0039
;  CURRENT APPLICATION NUMBER: US/09/503,505A
;  CURRENT FILING DATE:  2000-02-14
;  PRIOR APPLICATION NUMBER: JP 36367/1999
;  PRIOR FILING DATE: 1999-02-15
;  PRIOR APPLICATION NUMBER: JP 93777/1999
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 3
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Coriolus hirsutus
JS-09-503-505A-3

  Query Match             59.7%;  Score 43;  DB 4;  Length 213;
  Best Local Similarity   53.3%;  Pred. No. 39;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         14 GGGGVGKSALTIQFI 28


RESULT 13
JS-08-531-525-49
; Sequence 49, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  215 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Coprinus cinereus
US-08-531-525-49

  Query Match             59.7%;  Score 43;  DB 2;  Length 215;
  Best Local Similarity   53.3%;  Pred. No. 39;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         17 GGGGVGKSALTIQFI 31


RESULT 14
US-08-718-270A-49
; Sequence 49, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Coprinus cinereus
US-08-718-270A-49

Query Match 59.7%; Score 43; DB 2; Length 215;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         17 GGGGVGKSALTIQFI 31
```

RESULT 15
US-08-394-880B-2
; Sequence 2, Application US/08394880B
; Patent No. 5705352
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Multiple Drug Resistance Gene Of
; TITLE OF INVENTION: Aspergillus Fumigatus
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company/Patent Division
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: US
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,880B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant G., Thomas
; REGISTRATION NUMBER: 35784
; REFERENCE/DOCKET NUMBER: X-9682
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-2459
; TELEFAX: (317) 277-1917
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 791 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-394-880B-2

Query Match 59.7%; Score 43; DB 1; Length 791;
Best Local Similarity 43.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXIT 16
              |||:|:: |:|: ::|
Db        431 GGGMVQSGAITIGELT 446
```

Search completed: June 2, 2004, 18:13:56
Job time : 6.53488 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 17:58:08 ; Search time 19.1085 Seconds
(without alignments)
251.370 Million cell updates/sec

Title: US-10-092-367-138
Perfect score: 72
Sequence: 1 GGGXVRXSAXTLHXITP 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	100.0	17	6	ABJ38948	Abj38948 Linear Ga
2	72	100.0	17	6	ABJ38980	Abj38980 Linear Ga
3	66	91.7	95	6	ABJ38902	Abj38902 Conopepti
4	65	90.3	17	6	ABJ38850	Abj38850 Linear Ga
5	65	90.3	17	6	ABJ38903	Abj38903 Conopepti
6	64	88.9	17	6	ABJ38977	Abj38977 Linear Ga
7	62	86.1	17	6	ABJ38976	Abj38976 Linear Ga
8	58	80.6	95	6	ABJ38896	Abj38896 Conopepti
9	56	77.8	17	6	ABJ38945	Abj38945 Linear Ga
10	56	77.8	95	6	ABJ38894	Abj38894 Conopepti
11	54	75.0	17	6	ABJ38944	Abj38944 Linear Ga
12	49	68.1	17	6	ABJ38897	Abj38897 Conopepti
13	49	68.1	17	6	ABJ38847	Abj38847 Linear Ga
14	47	65.3	17	6	ABJ38846	Abj38846 Linear Ga
15	47	65.3	17	6	ABJ38895	Abj38895 Conopepti
16	47	65.3	1132	3	AAB53126	Aab53126 Macaca mu
17	46	63.9	17	6	ABJ38978	Abj38978 Linear Ga
18	46	63.9	97	6	ABJ38898	Abj38898 Conopepti
19	44	61.1	308	6	ADA32948	Ada32948 Acinetoba
20	44	61.1	339	2	AAW60076	Aaw60076 Escherich
21	44	61.1	805	6	ABP80438	Abp80438 N. gonorr
22	44	61.1	878	6	ADA83889	Ada83889 Human MUC
23	43	59.7	48	4	ABB15971	Abb15971 Human ner
24	43	59.7	140	4	ABG18776	Abg18776 Novel hum
25	43	59.7	195	4	ABB57799	Abb57799 Drosophil
26	43	59.7	203	2	AAW82591	Aaw82591 Human TC2
27	43	59.7	203	6	ABR41057	Abr41057 Human MAP
28	43	59.7	204	2	AAR77647	Aar77647 TC21 muta
29	43	59.7	204	6	ABR41056	Abr41056 Human MAP
30	43	59.7	210	2	AAY42695	Aay42695 Human R-R
31	43	59.7	213	3	AAB07940	Aab07940 Amino aci
32	43	59.7	218	5	AAU75736	Aau75736 Human rel
33	43	59.7	218	6	ABU62885	Abu62885 Ras-famil
34	43	59.7	252	4	ABG18778	Abg18778 Novel hum
35	43	59.7	288	4	AAG74576	Aag74576 Human col
36	43	59.7	371	6	ABM67376	Abm67376 Photorhab
37	43	59.7	589	3	AAB26937	Aab26937 Rice auxi
38	43	59.7	740	4	AAB99359	Aab99359 Human R-R
39	43	59.7	740	5	ABB06727	Abb06727 Human R-R
40	43	59.7	740	5	ABB06737	Abb06737 Human R-R
41	43	59.7	791	2	AAW01022	Aaw01022 Multiple
42	43	59.7	957	7	ADD47260	Add47260 Human Pro
43	43	59.7	957	7	ADE58049	Ade58049 Human Pro
44	43	59.7	957	7	ADD47264	Add47264 Human Pro
45	43	59.7	957	7	ADE58045	Ade58045 Human Pro

ALIGNMENTS

RESULT 1
ABJ38948
ID ABJ38948 standard; peptide; 17 AA.
XX
AC ABJ38948;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt5 SEQ ID No 138.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 4
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 7
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 10
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 14
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;

PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 43; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 72; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITP 17
||||||||||||||||
Db 1 GGGXVRXSAXTLHXITP 17

RESULT 2
ABJ38980
ID ABJ38980 standard; peptide; 17 AA.
XX
AC ABJ38980;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt5 SEQ ID No 170.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 44; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 72; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 7.3e-05;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITP 17

```
                  |||:||:||:|||:|||
b           1 GGGEVRESAETLHEITP 17


ESULT 3
ABJ38902
D   ABJ38902 standard; protein; 95 AA.
X
C   ABJ38902;
X
T   09-OCT-2003  (first entry)
X
E   Conopeptide conotoxin protein Bt5 SEQ ID No 73.
X
W   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
W   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
W   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
W   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
W   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
W   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
W   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
W   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
W   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
W   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
W   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
W   pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
W   parasitic worm.
X
S   Conus betulinus.
X
N   WO200272005-A2.
X
D   19-SEP-2002.
X
F   07-MAR-2002; 2002WO-US006863.
X
R   07-MAR-2001; 2001US-0273639P.
X
A   (UTAH ) UNIV UTAH RES FOUND.
A   (COGN-) COGNETIX INC.
X
I   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
I   Jones RM;
X
R   WPI; 2003-175000/17.
R   N-PSDB; ABT43476.
X
T   New conotoxins useful for treating e.g. neurologic disorders (e.g.
T   seizure associated with epilepsy or neurotoxic injury associated with
T   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
T   morphine tolerance).
X
S   Claim 5; Page 33; 113pp; English.
X
C   This invention relates to a novel isolated peptide consisting of
C   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
C   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
C   Sm1. The isolated conotoxin peptides are useful in methods for treating
C   or preventing disorders in which the pathophysiology involves excessive
C   excitation of nerve cells by excitatory amino acids or agonists of
C   heterogenous inotropic glutamate receptors or heterogenous B protein
C   coupled glutamate receptors; and for treating memory or cognitive
C   deficits, HIV infection, or ophthalmic indications comprising
C   administering to a patient a peptide above or its salt. Disorders include
C   neurological disorder or a psychiatric disorder, where the neurological
C   disorder is seizure associated with epilepsy or neurotoxic injury
C   associated with conditions of hypoxia, anoxia or ischaemia, including
C   neurotoxic injury associated with stroke, cerebrovascular accident, brain
C   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
C   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
C   disorder may also be a neurodegeneration associated with Alzheimer's
C   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
C   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
C   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
C   multi-infarct dementia, Binswanger dementia and neuronal damage
C   associated with uncontrolled seizures. The neurologic disorder is pain
C   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
C   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
C   barbiturate tolerance), dystonia (movement disorder), urinary
C   incontinence, muscle relaxation or sleep disorder. The psychiatric
C   disorder is anxiety, major depression, manic-depressive illness,
C   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
C   bipolar disorder, unipolar depression, dysthymia or seasonal effective
C   disorder. The conotoxin peptides are also useful for controlling
C   nematodes or parasitic worms by applying the peptides to the locus to be
C   protected. This sequence represents a linear gamma-carboxyglutamate rich
C   conotoxin protein of the invention
XX
SQ  Sequence 95 AA;

  Query Match             91.7%;  Score 66;  DB 6;  Length 95;
  Best Local Similarity   75.0%;  Pred. No. 0.0061;
  Matches   12;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              ||:||:||:|||:|||
Db         80 GGEVRESAETLHEITP 95


RESULT 4
ABJ38850
ID  ABJ38850 standard; peptide; 17 AA.
XX
AC  ABJ38850;
XX
DT  09-OCT-2003  (first entry)
XX
DE  Linear Gamma-carboxyglutamate rich conotoxin peptide SEQ ID No 6.
XX
KW  Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW  antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW  tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW  Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW  Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW  inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW  heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW  seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW  neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW  myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW  hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW  pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW  parasitic worm.
XX
OS  Conus betulinus.
XX
FH  Key             Location/Qualifiers
FT  Modified-site   4
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   7
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   10
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   14
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   17
FT                  /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
PF  07-MAR-2002; 2002WO-US006863.
XX
PR  07-MAR-2001; 2001US-0273639P.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
```

```
A   (COGN-) COGNETIX INC.
X
I   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
I   Jones RM;
X
R   WPI; 2003-175000/17.
X
T   New conotoxins useful for treating e.g. neurologic disorders (e.g.
T   seizure associated with epilepsy or neurotoxic injury associated with
T   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
T   morphine tolerance).
X
S   Claim 1; Page 48; 113pp; English.
X
C   This invention relates to a novel isolated peptide consisting of
C   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
C   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
C   Sm1. The isolated conotoxin peptides are useful in methods for treating
C   or preventing disorders in which the pathophysiology involves excessive
C   excitation of nerve cells by excitatory amino acids or agonists of
C   heterogenous inotropic glutamate receptors or heterogenous B protein
C   coupled glutamate receptors; and for treating memory or cognitive
C   deficits, HIV infection, or ophthalmic indications comprising
C   administering to a patient a peptide above or its salt. Disorders include
C   neurological disorder or a psychiatric disorder, where the neurological
C   disorder is seizure associated with epilepsy or neurotoxic injury
C   associated with conditions of hypoxia, anoxia or ischaemia, including
C   neurotoxic injury associated with stroke, cerebrovascular accident, brain
C   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
C   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
C   disorder may also be a neurodegeneration associated with Alzheimer's
'C  disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
'C  Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
:C  Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
:C  multi-infarct dementia, Binswanger dementia and neuronal damage
LC  associated with uncontrolled seizures. The neurologic disorder is pain
 C  (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
:C  addiction, morphine tolerance, opiate tolerance, opioid tolerance and
 C  barbiturate tolerance), dystonia (movement disorder), urinary
'C  incontinence, muscle relaxation or sleep disorder. The psychiatric
'C  disorder is anxiety, major depression, manic-depressive illness,
 C  obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
 C  bipolar disorder, unipolar depression, dysthymia or seasonal effective
'C  disorder. The conotoxin peptides are also useful for controlling
:C  nematodes or parasitic worms by applying the peptides to the locus to be
:C  protected. This sequence represents a linear gamma-carboxyglutamate rich
:C  conotoxin peptide of the invention
XX
SQ  Sequence 17 AA;

  Query Match             90.3%;  Score 65;  DB 6;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0011;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              ||||||||||||||||
Db          1 GGGXVRXSAXTLHXIT 16


RESULT 5
ABJ38903
ID  ABJ38903 standard; peptide; 17 AA.
XX
AC  ABJ38903;
XX
DT  09-OCT-2003  (first entry)
XX
DE  Conopeptide toxin peptide Bt5 SEQ ID No 74.
XX
KW  Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW  antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW  tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW  Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW  Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW  inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW  heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW  seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW  neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW  myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW  hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW  pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW  parasitic worm; toxin.
XX
OS  Conus betulinus.
XX
FH  Key             Location/Qualifiers
FT  Modified-site   4
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   7
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   10
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   14
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   17
FT                  /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
PF  07-MAR-2002; 2002WO-US006863.
XX
PR  07-MAR-2001; 2001US-0273639P.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
PA  (COGN-) COGNETIX INC.
XX
PI  Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI  Jones RM;
XX
DR  WPI; 2003-175000/17.
XX
PT  New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT  seizure associated with epilepsy or neurotoxic injury associated with
PT  hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT  morphine tolerance).
XX
PS  Example 7; Page 33; 113pp; English.
XX
CC  This invention relates to a novel isolated peptide consisting of
CC  conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC  Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC  Sm1. The isolated conotoxin peptides are useful in methods for treating
CC  or preventing disorders in which the pathophysiology involves excessive
CC  excitation of nerve cells by excitatory amino acids or agonists of
CC  heterogenous inotropic glutamate receptors or heterogenous B protein
CC  coupled glutamate receptors; and for treating memory or cognitive
CC  deficits, HIV infection, or ophthalmic indications comprising
CC  administering to a patient a peptide above or its salt. Disorders include
CC  neurological disorder or a psychiatric disorder, where the neurological
CC  disorder is seizure associated with epilepsy or neurotoxic injury
CC  associated with conditions of hypoxia, anoxia or ischaemia, including
CC  neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC  or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC  suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC  disorder may also be a neurodegeneration associated with Alzheimer's
CC  disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC  Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC  Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC  multi-infarct dementia, Binswanger dementia and neuronal damage
CC  associated with uncontrolled seizures. The neurologic disorder is pain
CC  (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC  addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC  barbiturate tolerance), dystonia (movement disorder), urinary
CC  incontinence, muscle relaxation or sleep disorder. The psychiatric
```

```
C   disorder is anxiety, major depression, manic-depressive illness,
C   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
C   bipolar disorder, unipolar depression, dysthymia or seasonal effective
C   disorder. The conotoxin peptides are also useful for controlling
C   nematodes or parasitic worms by applying the peptides to the locus to be
C   protected. This sequence represents a toxin sequence of a linear gamma-
C   carboxyglutamate rich conotoxin peptide of the invention
X
Q   Sequence 17 AA;

 Query Match             90.3%;  Score 65;  DB 6;  Length 17;
 Best Local Similarity   100.0%;  Pred. No. 0.0011;
 Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              ||||||||||||||||
Db          1 GGGXVRXSAXTLHXIT 16


RESULT 6
ABJ38977
 D   ABJ38977 standard; peptide; 17 AA.
XX
AC   ABJ38977;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide Bt2 SEQ ID No 167.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Example 7; Page 44; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             88.9%;  Score 64;  DB 6;  Length 17;
  Best Local Similarity   70.6%;  Pred. No. 0.0016;
  Matches   12;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || :||:||:|||:|||
Db          1 GGEEVRESAETLHEITP 17


RESULT 7
ABJ38976
ID   ABJ38976 standard; peptide; 17 AA.
XX
AC   ABJ38976;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide Bt1 SEQ ID No 166.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
```

```
X
A   (UTAH ) UNIV UTAH RES FOUND.
A   (COGN-) COGNETIX INC.
X
I   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
I   Jones RM;
X
R   WPI; 2003-175000/17.
X
T   New conotoxins useful for treating e.g. neurologic disorders (e.g.
T   seizure associated with epilepsy or neurotoxic injury associated with
T   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
T   morphine tolerance).
X
S   Example 7; Page 44; 113pp; English.
X
C   This invention relates to a novel isolated peptide consisting of
C   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
C   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
C   Sm1. The isolated conotoxin peptides are useful in methods for treating
C   or preventing disorders in which the pathophysiology involves excessive
C   excitation of nerve cells by excitatory amino acids or agonists of
C   heterogenous inotropic glutamate receptors or heterogenous B protein
C   coupled glutamate receptors; and for treating memory or cognitive
C   deficits, HIV infection, or ophthalmic indications comprising
C   administering to a patient a peptide above or its salt. Disorders include
C   neurological disorder or a psychiatric disorder, where the neurological
'C  disorder is seizure associated with epilepsy or neurotoxic injury
'C  associated with conditions of hypoxia, anoxia or ischaemia, including
'C  neurotoxic injury associated with stroke, cerebrovascular accident, brain
'C  or spinal cord trauma, myocardial infarct, physical trauma, drownings,
'C  suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
'C  disorder may also be a neurodegeneration associated with Alzheimer's
'C  disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
C   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
C   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
C   multi-infarct dementia, Binswanger dementia and neuronal damage
C   associated with uncontrolled seizures. The neurologic disorder is pain
C   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
C   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
C   barbiturate tolerance), dystonia (movement disorder), urinary
C   incontinence, muscle relaxation or sleep disorder. The psychiatric
C   disorder is anxiety, major depression, manic-depressive illness,
'C  obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
'C  bipolar disorder, unipolar depression, dysthymia or seasonal effective
'C  disorder. The conotoxin peptides are also useful for controlling
'C  nematodes or parasitic worms by applying the peptides to the locus to be
'C  protected. This sequence represents a linear gamma-carboxyglutamate rich
'C  conotoxin peptide of the invention
XX
SQ  Sequence 17 AA;

  Query Match             86.1%;  Score 62;  DB 6;  Length 17;
  Best Local Similarity   64.7%;  Pred. No. 0.0034;
  Matches   11;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || :||:||:|||::||
Db          1 GGEEVRESAETLHELTP 17


RESULT 8
ABJ38896
ID  ABJ38896 standard; protein; 95 AA.
XX
AC  ABJ38896;
XX
DT  09-OCT-2003  (first entry)
XX
DE  Conopeptide conotoxin protein Bt2 SEQ ID No 64.
XX
KW  Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW  antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW  tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW  Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW  Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW  inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW  heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW  seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW  neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW  myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW  hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW  pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW  parasitic worm.
XX
OS  Conus betulinus.
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
PF  07-MAR-2002; 2002WO-US006863.
XX
PR  07-MAR-2001; 2001US-0273639P.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
PA  (COGN-) COGNETIX INC.
XX
PI  Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI  Jones RM;
XX
DR  WPI; 2003-175000/17.
DR  N-PSDB; ABT43473.
XX
PT  New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT  seizure associated with epilepsy or neurotoxic injury associated with
PT  hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT  morphine tolerance).
XX
PS  Claim 5; Page 32; 113pp; English.
XX
CC  This invention relates to a novel isolated peptide consisting of
CC  conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC  Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC  Sm1. The isolated conotoxin peptides are useful in methods for treating
CC  or preventing disorders in which the pathophysiology involves excessive
CC  excitation of nerve cells by excitatory amino acids or agonists of
CC  heterogenous inotropic glutamate receptors or heterogenous B protein
CC  coupled glutamate receptors; and for treating memory or cognitive
CC  deficits, HIV infection, or ophthalmic indications comprising
CC  administering to a patient a peptide above or its salt. Disorders include
CC  neurological disorder or a psychiatric disorder, where the neurological
CC  disorder is seizure associated with epilepsy or neurotoxic injury
CC  associated with conditions of hypoxia, anoxia or ischaemia, including
CC  neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC  or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC  suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC  disorder may also be a neurodegeneration associated with Alzheimer's
CC  disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC  Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC  Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC  multi-infarct dementia, Binswanger dementia and neuronal damage
CC  associated with uncontrolled seizures. The neurologic disorder is pain
CC  (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC  addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC  barbiturate tolerance), dystonia (movement disorder), urinary
CC  incontinence, muscle relaxation or sleep disorder. The psychiatric
CC  disorder is anxiety, major depression, manic-depressive illness,
CC  obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC  bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC  disorder. The conotoxin peptides are also useful for controlling
CC  nematodes or parasitic worms by applying the peptides to the locus to be
CC  protected. This sequence represents a linear gamma-carboxyglutamate rich
CC  conotoxin protein of the invention
XX
SQ  Sequence 95 AA;
```

```
 Query Match             80.6%;  Score 58;  DB 6;  Length 95;
 Best Local Similarity   68.8%;  Pred. No. 0.13;
 Matches   11;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Y          2 GGXVRXSAXTLHXITP 17
             | :||:||:|||:|||
D         80 GEEVRESAETLHEITP 95


ESULT 9
BJ38945
D   ABJ38945 standard; peptide; 17 AA.
X
C   ABJ38945;
X
T   09-OCT-2003  (first entry)
X
E   Linear Gamma-carboxyglutamate rich conotoxin peptide Bt2 SEQ ID No 135.
X
W   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
W   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
W   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
W   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
W   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
W   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
W   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
W   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
W   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
W   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
W   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
W   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
W   parasitic worm.
X
S   Conus betulinus.
X
H   Key             Location/Qualifiers
T   Modified-site   3
T                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
T   Modified-site   4
T                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
T   Modified-site   7
T                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
T   Modified-site   10
T                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
T   Modified-site   14
T                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
PF  07-MAR-2002; 2002WO-US006863.
XX
PR  07-MAR-2001; 2001US-0273639P.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
PA  (COGN-) COGNETIX INC.
XX
PI  Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI  Jones RM;
XX
DR  WPI; 2003-175000/17.
XX
PT  New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT  seizure associated with epilepsy or neurotoxic injury associated with
PT  hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT  morphine tolerance).
XX
PS  Example 7; Page 43; 113pp; English.
XX
CC  This invention relates to a novel isolated peptide consisting of
CC  conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC  Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC  Sm1. The isolated conotoxin peptides are useful in methods for treating
CC  or preventing disorders in which the pathophysiology involves excessive
CC  excitation of nerve cells by excitatory amino acids or agonists of
CC  heterogenous inotropic glutamate receptors or heterogenous B protein
CC  coupled glutamate receptors; and for treating memory or cognitive
CC  deficits, HIV infection, or ophthalmic indications comprising
CC  administering to a patient a peptide above or its salt. Disorders include
CC  neurological disorder or a psychiatric disorder, where the neurological
CC  disorder is seizure associated with epilepsy or neurotoxic injury
CC  associated with conditions of hypoxia, anoxia or ischaemia, including
CC  neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC  or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC  suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC  disorder may also be a neurodegeneration associated with Alzheimer's
CC  disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC  Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC  Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC  multi-infarct dementia, Binswanger dementia and neuronal damage
CC  associated with uncontrolled seizures. The neurologic disorder is pain
CC  (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC  addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC  barbiturate tolerance), dystonia (movement disorder), urinary
CC  incontinence, muscle relaxation or sleep disorder. The psychiatric
CC  disorder is anxiety, major depression, manic-depressive illness,
CC  obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC  bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC  disorder. The conotoxin peptides are also useful for controlling
CC  nematodes or parasitic worms by applying the peptides to the locus to be
CC  protected. This sequence represents a linear gamma-carboxyglutamate rich
CC  conotoxin peptide of the invention
XX
SQ  Sequence 17 AA;

  Query Match             77.8%;  Score 56;  DB 6;  Length 17;
  Best Local Similarity   94.1%;  Pred. No. 0.034;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || |||||||||||||||
Db          1 GGXXVRXSAXTLHXITP 17


RESULT 10
ABJ38894
ID  ABJ38894 standard; protein; 95 AA.
XX
AC  ABJ38894;
XX
DT  09-OCT-2003  (first entry)
XX
DE  Conopeptide conotoxin protein Bt1 SEQ ID No 61.
XX
KW  Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW  antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW  tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW  Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW  Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW  inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW  heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW  seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW  neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW  myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW  hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW  pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW  parasitic worm.
XX
OS  Conus betulinus.
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
```

F 07-MAR-2002; 2002WO-US006863.
X
R 07-MAR-2001; 2001US-0273639P.
X
A (UTAH) UNIV UTAH RES FOUND.
A (COGN-) COGNETIX INC.
X
I Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
I Jones RM;
X
R WPI; 2003-175000/17.
R N-PSDB; ABT43472.
X
T New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 31; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 95 AA;

Query Match 77.8%; Score 56; DB 6; Length 95;
Best Local Similarity 62.5%; Pred. No. 0.28;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GGXVRXSAXTLHXITP 17
              | :||:||:|||::||
Db         80 GEEVRESAETLHELTP 95
```

RESULT 11
ABJ38944
ID ABJ38944 standard; peptide; 17 AA.
XX
AC ABJ38944;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt1 SEQ ID No 134.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX

FH	Key	Location/Qualifiers
FT	Modified-site	3
FT		/note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT	Modified-site	4
FT		/note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT	Modified-site	7
FT		/note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT	Modified-site	10
FT		/note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT	Modified-site	14
FT		/note= "Residue is optionally Glu or gamma-carboxy-Glu"

XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 43; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,

```
C   multi-infarct dementia, Binswanger dementia and neuronal damage
C   associated with uncontrolled seizures. The neurologic disorder is pain
C   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
C   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
C   barbiturate tolerance), dystonia (movement disorder), urinary
C   incontinence, muscle relaxation or sleep disorder. The psychiatric
C   disorder is anxiety, major depression, manic-depressive illness,
C   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
C   bipolar disorder, unipolar depression, dysthymia or seasonal effective
C   disorder. The conotoxin peptides are also useful for controlling
C   nematodes or parasitic worms by applying the peptides to the locus to be
C   protected. This sequence represents a linear gamma-carboxyglutamate rich
C   conotoxin peptide of the invention
X
Q   Sequence 17 AA;

 Query Match             75.0%;  Score 54;  DB 6;  Length 17;
 Best Local Similarity   88.2%;  Pred. No. 0.073;
 Matches   15;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

y          1 GGGXVRXSAXTLHXITP 17
             || |||||||||||:||
b          1 GGXXVRXSAXTLHXLTP 17


RESULT 12
ABJ38897
ID   ABJ38897 standard; peptide; 17 AA.
XX
AC   ABJ38897;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide toxin peptide Bt2 SEQ ID No 65.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm; toxin.
XX
OS   Conus betulinus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   3
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   4
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   7
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   10
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   14
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   17
FT                   /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Example 7; Page 32; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a toxin sequence of a linear gamma-
CC   carboxyglutamate rich conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             68.1%;  Score 49;  DB 6;  Length 17;
  Best Local Similarity   93.8%;  Pred. No. 0.5;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||||
Db          1 GGXXVRXSAXTLHXIT 16


RESULT 13
ABJ38847
ID   ABJ38847 standard; peptide; 17 AA.
XX
AC   ABJ38847;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide SEQ ID No 3.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
```

```
W   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
W   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
W   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
W   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
W   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
W   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
W   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
W   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
W   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
W   parasitic worm.
XX
OS  Conus betulinus.
XX
FH  Key             Location/Qualifiers
FT  Modified-site   4
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   7
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   10
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   14
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   17
FT                  /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
PF  07-MAR-2002; 2002WO-US006863.
XX
PR  07-MAR-2001; 2001US-0273639P.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
PA  (COGN-) COGNETIX INC.
XX
PI  Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI  Jones RM;
XX
DR  WPI; 2003-175000/17.
XX
PT  New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT  seizure associated with epilepsy or neurotoxic injury associated with
PT  hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT  morphine tolerance).
XX
PS  Claim 1; Page 48; 113pp; English.
XX
CC  This invention relates to a novel isolated peptide consisting of
CC  conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC  Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC  Sm1. The isolated conotoxin peptides are useful in methods for treating
CC  or preventing disorders in which the pathophysiology involves excessive
CC  excitation of nerve cells by excitatory amino acids or agonists of
CC  heterogenous inotropic glutamate receptors or heterogenous B protein
CC  coupled glutamate receptors; and for treating memory or cognitive
CC  deficits, HIV infection, or ophthalmic indications comprising
CC  administering to a patient a peptide above or its salt. Disorders include
CC  neurological disorder or a psychiatric disorder, where the neurological
CC  disorder is seizure associated with epilepsy or neurotoxic injury
CC  associated with conditions of hypoxia, anoxia or ischaemia, including
CC  neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC  or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC  suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC  disorder may also be a neurodegeneration associated with Alzheimer's
CC  disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC  Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC  Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC  multi-infarct dementia, Binswanger dementia and neuronal damage
CC  associated with uncontrolled seizures. The neurologic disorder is pain
CC  (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC  addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC  barbiturate tolerance), dystonia (movement disorder), urinary
CC  incontinence, muscle relaxation or sleep disorder. The psychiatric
CC  disorder is anxiety, major depression, manic-depressive illness,
CC  obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC  bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC  disorder. The conotoxin peptides are also useful for controlling
CC  nematodes or parasitic worms by applying the peptides to the locus to be
CC  protected. This sequence represents a linear gamma-carboxyglutamate rich
CC  conotoxin peptide of the invention
XX
SQ  Sequence 17 AA;

  Query Match             68.1%;  Score 49;  DB 6;  Length 17;
  Best Local Similarity   93.8%;  Pred. No. 0.5;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||||
Db          1 GGXXVRXSAXTLHXIT 16


RESULT 14
ABJ38846
ID  ABJ38846 standard; peptide; 17 AA.
XX
AC  ABJ38846;
XX
DT  09-OCT-2003  (first entry)
XX
DE  Linear Gamma-carboxyglutamate rich conotoxin peptide SEQ ID No 2.
XX
KW  Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW  antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW  tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW  Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW  Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW  inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW  heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW  seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW  neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW  myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW  hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW  pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW  parasitic worm.
XX
OS  Conus betulinus.
XX
FH  Key             Location/Qualifiers
FT  Modified-site   3
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   4
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   7
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   10
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   14
FT                  /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT  Modified-site   17
FT                  /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN  WO200272005-A2.
XX
PD  19-SEP-2002.
XX
PF  07-MAR-2002; 2002WO-US006863.
XX
PR  07-MAR-2001; 2001US-0273639P.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
PA  (COGN-) COGNETIX INC.
XX
PI  Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI  Jones RM;
```

```
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Claim 1; Page 48; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             65.3%;  Score 47;  DB 6;  Length 17;
  Best Local Similarity   87.5%;  Pred. No. 1.1;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||:|
Db          1 GGXXVRXSAXTLHXLT 16


RESULT 15
ABJ38895
ID   ABJ38895 standard; peptide; 17 AA.
XX
AC   ABJ38895;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide toxin peptide Bt1 SEQ ID No 62.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm; toxin.
XX
OS   Conus betulinus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   3
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   4
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   7
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   10
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   14
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   17
FT                   /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Example 7; Page 31; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
```

C bipolar disorder, unipolar depression, dysthymia or seasonal effective
C disorder. The conotoxin peptides are also useful for controlling
C nematodes or parasitic worms by applying the peptides to the locus to be
C protected. This sequence represents a toxin sequence of a linear gamma-
C carboxyglutamate rich conotoxin peptide of the invention
XX
Q Sequence 17 AA;

Query Match 65.3%; Score 47; DB 6; Length 17;
Best Local Similarity 87.5%; Pred. No. 1.1;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXIT 16
|| |||||||||||:|
b 1 GGXXVRXSAXTLHXLT 16

earch completed: June 2, 2004, 18:09:45
Job time : 20.1085 secs

GenCore version 5.1.6

M protein - protein search, using sw model

un on: June 2, 2004, 18:13:14 ; Search time 14.3643 Seconds
(without alignments)
332.960 Million cell updates/sec

itle: US-10-092-367-138
erfect score: 72
equence: 1 GGGXVRXSAXTLHXITP 17

coring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

earched: 1155919 seqs, 281338677 residues

'otal number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	100.0	17	12	US-10-092-367-138	Sequence 138, App
2	72	100.0	17	12	US-10-092-367-170	Sequence 170, App
3	66	91.7	95	12	US-10-092-367-73	Sequence 73, Appl
4	65	90.3	17	12	US-10-092-367-6	Sequence 6, Appli
5	65	90.3	17	12	US-10-092-367-74	Sequence 74, Appl
6	64	88.9	17	12	US-10-092-367-167	Sequence 167, App
7	62	86.1	17	12	US-10-092-367-166	Sequence 166, App
8	58	80.6	95	12	US-10-092-367-64	Sequence 64, Appl
9	56	77.8	17	12	US-10-092-367-135	Sequence 135, App
10	56	77.8	95	12	US-10-092-367-61	Sequence 61, Appl
11	54	75.0	17	12	US-10-092-367-134	Sequence 134, App
12	49	68.1	17	12	US-10-092-367-3	Sequence 3, Appli
13	49	68.1	17	12	US-10-092-367-65	Sequence 65, Appl
14	47	65.3	17	12	US-10-092-367-2	Sequence 2, Appli
15	47	65.3	17	12	US-10-092-367-62	Sequence 62, Appl
16	46	63.9	17	12	US-10-092-367-168	Sequence 168, App
17	46	63.9	97	12	US-10-092-367-67	Sequence 67, Appl
18	45	62.5	333	12	US-10-424-599-244796	Sequence 244796,
19	44	61.1	164	12	US-10-424-599-207206	Sequence 207206,
20	44	61.1	878	14	US-10-157-031-171	Sequence 171, App
21	43	59.7	97	12	US-10-425-114-68008	Sequence 68008, A
22	43	59.7	102	12	US-10-424-599-200067	Sequence 200067,
23	43	59.7	136	12	US-10-425-114-56785	Sequence 56785, A
24	43	59.7	137	12	US-10-425-114-67405	Sequence 67405, A
25	43	59.7	192	12	US-10-424-599-282377	Sequence 282377,
26	43	59.7	203	14	US-10-197-666A-84	Sequence 84, Appl
27	43	59.7	204	14	US-10-197-666A-82	Sequence 82, Appl
28	43	59.7	204	16	US-10-408-765A-1241	Sequence 1241, Ap
29	43	59.7	218	10	US-09-873-546-14	Sequence 14, Appl
30	43	59.7	218	13	US-10-067-813-17	Sequence 17, Appl
31	43	59.7	218	16	US-10-408-765A-690	Sequence 690, App
32	43	59.7	231	15	US-10-369-493-5603	Sequence 5603, Ap
33	43	59.7	274	15	US-10-369-493-15531	Sequence 15531, A
34	43	59.7	275	15	US-10-369-493-15902	Sequence 15902, A
35	43	59.7	275	15	US-10-369-493-16274	Sequence 16274, A
36	43	59.7	275	15	US-10-369-493-17567	Sequence 17567, A
37	43	59.7	287	15	US-10-369-493-9354	Sequence 9354, Ap
38	43	59.7	288	14	US-10-106-698-5350	Sequence 5350, Ap
39	43	59.7	309	9	US-09-801-368-286	Sequence 286, App
40	43	59.7	309	15	US-10-369-493-22369	Sequence 22369, A
41	43	59.7	557	15	US-10-369-493-20460	Sequence 20460, A
42	43	59.7	740	12	US-10-344-404-23	Sequence 23, Appl
43	43	59.7	957	10	US-09-840-746-19	Sequence 19, Appl
44	42.5	59.0	1684	12	US-10-276-774-2398	Sequence 2398, Ap
45	42.5	59.0	1727	16	US-10-408-765A-1813	Sequence 1813, Ap

ALIGNMENTS

RESULT 1
US-10-092-367-138
; Sequence 138, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 138
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Xaa is Glu or gamma-carboxy-Glu
US-10-092-367-138

Query Match 100.0%; Score 72; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITP 17
|||||||||||||||||
Db 1 GGGXVRXSAXTLHXITP 17

```
RESULT 2
US-10-092-367-170
, Sequence 170, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
,  APPLICANT:  Olivera, Baldomero M
,  APPLICANT:  McIntosh, J. Michael
,  APPLICANT:  Garrett, James E.
,  APPLICANT:  Walker, Craig S.
,  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
,  FILE REFERENCE: 2314-224-II
,  CURRENT APPLICATION NUMBER: US/10/092,367
,  CURRENT FILING DATE:  2002-03-07
,  PRIOR APPLICATION NUMBER: US 60/273,639
,  PRIOR FILING DATE: 2001-03-07
,  NUMBER OF SEQ ID NOS: 196
,  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
US-10-092-367-170

  Query Match             100.0%;  Score 72;  DB 12;  Length 17;
  Best Local Similarity    76.5%;  Pred. No. 0.00036;
  Matches   13;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              |||:||:||:|||:|||
Db          1 GGGEVRESAETLHEITP 17


RESULT 3
US-10-092-367-73
, Sequence 73, Application US/10092367
, Publication No. US20030065138A1
, GENERAL INFORMATION:
,  APPLICANT: University of Utah Research Foundation
,  APPLICANT:  Cognetix, Inc.
,  APPLICANT:  Olivera, Baldomero M
,  APPLICANT:  McIntosh, J. Michael
,  APPLICANT:  Garrett, James E.
,  APPLICANT:  Walker, Craig S.
,  APPLICANT:  Watkins, Maren
,  APPLICANT:  Jones, Robert M.
,  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
,  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
;   LENGTH: 95
;   TYPE: PRT
;   ORGANISM: Conus betulinus
US-10-092-367-73

  Query Match             91.7%;  Score 66;  DB 12;  Length 95;
  Best Local Similarity   75.0%;  Pred. No. 0.022;
  Matches   12;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              ||:||:||:|||:|||
Db         80 GGEVRESAETLHEITP 95


RESULT 4
US-10-092-367-6
; Sequence 6, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; Xaa
;   OTHER INFORMATION: at residue 17 is Pro or hydroxy-Pro
US-10-092-367-6

  Query Match             90.3%;  Score 65;  DB 12;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0043;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              ||||||||||||||||
Db          1 GGGXVRXSAXTLHXIT 16


RESULT 5
US-10-092-367-74
; Sequence 74, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; Xaa
;   OTHER INFORMATION: at residue 17 is Pro or hydroxy-Pro
US-10-092-367-74
```

```
 Query Match             90.3%;  Score 65;  DB 12;  Length 17;
 Best Local Similarity   100.0%;  Pred. No. 0.0043;
 Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y          1 GGGXVRXSAXTLHXIT 16
             ||||||||||||||||
b          1 GGGXVRXSAXTLHXIT 16


ESULT 6
S-10-092-367-167
 Sequence 167, Application US/10092367
 Publication No. US20030065138A1
 GENERAL INFORMATION:
  APPLICANT: University of Utah Research Foundation
  APPLICANT:  Cognetix, Inc.
  APPLICANT:  Olivera, Baldomero M
  APPLICANT:  McIntosh, J. Michael
  APPLICANT:  Garrett, James E.
  APPLICANT:  Walker, Craig S.
  APPLICANT:  Watkins, Maren
  APPLICANT:  Jones, Robert M.
  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
  FILE REFERENCE: 2314-224-II
  CURRENT APPLICATION NUMBER: US/10/092,367
  CURRENT FILING DATE:  2002-03-07
  PRIOR APPLICATION NUMBER: US 60/273,639
  PRIOR FILING DATE: 2001-03-07
  NUMBER OF SEQ ID NOS: 196
  SOFTWARE: PatentIn version 3.0
 SEQ ID NO 167
   LENGTH: 17
   TYPE: PRT
   ORGANISM: Conus betulinus
JS-10-092-367-167

 Query Match             88.9%;  Score 64;  DB 12;  Length 17;
 Best Local Similarity   70.6%;  Pred. No. 0.0062;
 Matches   12;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || :||:||:|||:|||
Db          1 GGEEVRESAETLHEITP 17


RESULT 7
JS-10-092-367-166
 Sequence 166, Application US/10092367
 Publication No. US20030065138A1
 GENERAL INFORMATION:
  APPLICANT: University of Utah Research Foundation
  APPLICANT:  Cognetix, Inc.
  APPLICANT:  Olivera, Baldomero M
  APPLICANT:  McIntosh, J. Michael
  APPLICANT:  Garrett, James E.
  APPLICANT:  Walker, Craig S.
  APPLICANT:  Watkins, Maren
  APPLICANT:  Jones, Robert M.
  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
  FILE REFERENCE: 2314-224-II
  CURRENT APPLICATION NUMBER: US/10/092,367
  CURRENT FILING DATE:  2002-03-07
  PRIOR APPLICATION NUMBER: US 60/273,639
  PRIOR FILING DATE: 2001-03-07
  NUMBER OF SEQ ID NOS: 196
  SOFTWARE: PatentIn version 3.0
 SEQ ID NO 166
   LENGTH: 17
   TYPE: PRT
   ORGANISM: Conus betulinus
JS-10-092-367-166

 Query Match             86.1%;  Score 62;  DB 12;  Length 17;
 Best Local Similarity   64.7%;  Pred. No. 0.013;
 Matches   11;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || :||:||:|||::||
Db          1 GGEEVRESAETLHELTP 17


RESULT 8
US-10-092-367-64
; Sequence 64, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 64
;   LENGTH: 95
;   TYPE: PRT
;   ORGANISM: Conus betulinus
US-10-092-367-64

 Query Match             80.6%;  Score 58;  DB 12;  Length 95;
 Best Local Similarity   68.8%;  Pred. No. 0.37;
 Matches   11;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              | :||:||:|||:|||
Db         80 GEEVRESAETLHEITP 95


RESULT 9
US-10-092-367-135
; Sequence 135, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 135
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
```

```
   OTHER INFORMATION: Xaa is Glu or gamma-carboxy-Glu
US-10-092-367-135

  Query Match             77.8%;  Score 56;  DB 12;  Length 17;
  Best Local Similarity   94.1%;  Pred. No. 0.11;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || |||||||||||||||
Db          1 GGXXVRXSAXTLHXITP 17


RESULT 10
US-10-092-367-61
  Sequence 61, Application US/10092367
  Publication No. US20030065138A1
  GENERAL INFORMATION:
   APPLICANT: University of Utah Research Foundation
   APPLICANT:  Cognetix, Inc.
   APPLICANT:  Olivera, Baldomero M
   APPLICANT:  McIntosh, J. Michael
   APPLICANT:  Garrett, James E.
   APPLICANT:  Walker, Craig S.
   APPLICANT:  Watkins, Maren
   APPLICANT:  Jones, Robert M.
   TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
   FILE REFERENCE: 2314-224-II
   CURRENT APPLICATION NUMBER: US/10/092,367
   CURRENT FILING DATE:  2002-03-07
   PRIOR APPLICATION NUMBER: US 60/273,639
   PRIOR FILING DATE: 2001-03-07
   NUMBER OF SEQ ID NOS: 196
   SOFTWARE: PatentIn version 3.0
  SEQ ID NO 61
     LENGTH: 95
     TYPE: PRT
     ORGANISM: Conus betulinus
US-10-092-367-61

  Query Match             77.8%;  Score 56;  DB 12;  Length 95;
  Best Local Similarity   62.5%;  Pred. No. 0.76;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              | :||:||:|||::||
Db         80 GEEVRESAETLHELTP 95


RESULT 11
US-10-092-367-134
; Sequence 134, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 134
;    LENGTH: 17
;    TYPE: PRT
;    ORGANISM: Conus betulinus
;    FEATURE:
;    NAME/KEY: PEPTIDE
;    LOCATION: (1)..(17)
;    OTHER INFORMATION: Xaa is Glu or gamma-carboxy-Glu
US-10-092-367-134

  Query Match             75.0%;  Score 54;  DB 12;  Length 17;
  Best Local Similarity   88.2%;  Pred. No. 0.21;
  Matches   15;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || |||||||||||:||
Db          1 GGXXVRXSAXTLHXLTP 17


RESULT 12
US-10-092-367-3
; Sequence 3, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;    LENGTH: 17
;    TYPE: PRT
;    ORGANISM: Conus betulinus
;    FEATURE:
;    NAME/KEY: PEPTIDE
;    LOCATION: (1)..(17)
;    OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
;    OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-3

  Query Match             68.1%;  Score 49;  DB 12;  Length 17;
  Best Local Similarity   93.8%;  Pred. No. 1.3;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||||
Db          1 GGXXVRXSAXTLHXIT 16


RESULT 13
US-10-092-367-65
; Sequence 65, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
```

```
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 65
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
;   OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-65

  Query Match             68.1%;  Score 49;  DB 12;  Length 17;
  Best Local Similarity   93.8%;  Pred. No. 1.3;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||||
Db          1 GGXXVRXSAXTLHXIT 16


RESULT 14
US-10-092-367-2
; Sequence 2, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
;   OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-2

  Query Match             65.3%;  Score 47;  DB 12;  Length 17;
  Best Local Similarity   87.5%;  Pred. No. 2.6;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||:|
Db          1 GGXXVRXSAXTLHXLT 16


RESULT 15
US-10-092-367-62
; Sequence 62, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
;   OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-62

  Query Match             65.3%;  Score 47;  DB 12;  Length 17;
  Best Local Similarity   87.5%;  Pred. No. 2.6;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || |||||||||||:|
Db          1 GGXXVRXSAXTLHXLT 16


Search completed: June 2, 2004, 18:15:58
Job time : 14.3643 secs
```

GenCore version 5.1.6

M protein - protein search, using sw model

Run on: June 2, 2004, 18:10:29 ; Search time 4.6124 Seconds
(without alignments)
354.534 Million cell updates/sec

Title: US-10-092-367-138
Perfect score: 72
Sequence: 1 GGGXVRXSAXTLHXITP 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	66.7	1417	2	H83132	probable sensor/re
2	45	62.5	343	2	G91161	RNA 3'-terminal ph
3	45	62.5	403	2	T40473	hypothetical prote
4	44	61.1	146	2	G65137	hypothetical 15.4
5	44	61.1	209	2	S13179	transforming prote
6	44	61.1	265	2	S40209	tubulin gamma chai
7	44	61.1	319	2	S76960	hypothetical prote
8	43	59.7	186	1	TVDORS	transforming prote
9	43	59.7	189	1	TVDORA	transforming prote
10	43	59.7	189	2	S33796	ras protein homolo
11	43	59.7	191	2	JC6328	Ras2 protein - sli
12	43	59.7	191	2	S58220	transforming prote
13	43	59.7	192	2	S55022	transforming prote
14	43	59.7	192	2	S32042	GTP-binding protei
15	43	59.7	193	2	S38362	Ppras2 protein - s
16	43	59.7	195	1	TVFFR	transforming prote
17	43	59.7	203	1	TVHUC2	GTP-binding protei
18	43	59.7	203	2	A36365	transforming prote
19	43	59.7	206	2	C36365	transforming prote
20	43	59.7	215	2	JN0562	hypothetical 24K p
21	43	59.7	217	1	TVWYRS	transforming prote
22	43	59.7	217	2	H70631	hypothetical prote
23	43	59.7	218	1	TVHURR	transforming prote
24	43	59.7	231	2	T32953	hypothetical prote
25	43	59.7	275	2	C82717	50S ribosomal prot
26	43	59.7	309	1	TVBYR1	GTP-binding protei
27	43	59.7	792	2	A84308	chloride channel [
28	43	59.7	814	2	T30950	hypothetical prote
29	43	59.7	2712	2	T30949	hypothetical prote
30	42	58.3	118	2	AH3454	hypothetical prote
31	42	58.3	345	2	C90416	hypothetical prote
32	42	58.3	433	2	T47750	N2,N2-dimethylguan
33	42	58.3	1145	2	S55600	single-stranded DN
34	42	58.3	1166	2	T13958	synGAP-b1 protein
35	42	58.3	1249	2	T14270	Ras-GTPase activat
36	42	58.3	1293	2	T14259	ras GTPase-activat
37	41	56.9	81	2	H84344	hypothetical prote
38	41	56.9	273	2	T35153	hypothetical prote
39	41	56.9	309	2	T25800	C2H2-type zinc fin
40	41	56.9	390	2	F96723	hypothetical prote
41	41	56.9	400	2	C75336	serine proteinase,
42	41	56.9	418	2	E69014	hypothetical prote
43	41	56.9	544	2	E75569	probable aminotran
44	41	56.9	650	2	T23175	hypothetical prote
45	41	56.9	654	2	A57785	finger protein ZNF

ALIGNMENTS

RESULT 1
H83132
probable sensor/response regulator hybrid PA4112 [imported] - Pseudomonas aeruginosa (str
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: H83132
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83132
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1417 <STO>
A;Cross-references: GB:AE004827; GB:AE004091; NID:g9950306; PIDN:AAG07499.1; GSPDB:GN0013
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4112

Query Match 66.7%; Score 48; DB 2; Length 1417;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              | |:|::||:|||:|
Db       1336 GEGDVQGSAATLHTI 1350
```

RESULT 2
G91161
RNA 3'-terminal phosphate cyclase [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: G91161
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G91161
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-343 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37686.1; PID:g13363737; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4263
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0025

Query Match 62.5%; Score 45; DB 2; Length 343;

```
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :||::| :||
Db         15 GGGQIMRSALSLSMIT 30


RESULT 3
T40473
hypothetical protein SPBC4B4.01c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C;Accession: T40473
R;Beck, A.; Reinhardt, R.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z21932
A;Accession: T40473
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-403 <BEC>
A;Cross-references: EMBL:AL023706; PIDN:CAA19281.1; GSPDB:GN00067; SPDB:SPBC4B4.01c
A;Experimental source: strain 972h-; cosmid c4B4
C;Genetics:
A;Gene: SPDB:SPBC4B4.01c
A;Map position: 2
A;Introns: 36/1; 65/3
C;Superfamily: Saccharomyces hypothetical protein YDR531w

  Query Match             62.5%;  Score 45;  DB 2;  Length 403;
  Best Local Similarity   37.5%;  Pred. No. 26;
  Matches    6;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              || ::|:  :|:|::|
Db        343 GGSFIRNHVQTMHTLT 358


RESULT 4
G65137
hypothetical 15.4 kD protein in malT-glpR intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: G65137
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65137
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-146 <BLAT>
A;Cross-references: GB:AE000418; GB:U00096; NID:g2367222; PIDN:AAC76445.1; PID:g1789826;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yhgK

  Query Match             61.1%;  Score 44;  DB 2;  Length 146;
  Best Local Similarity   50.0%;  Pred. No. 12;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :||::| :||
Db         15 GGGQILRSALSLSMIT 30


RESULT 5
S13179
transforming protein (ras) - Geodia cydonium
C;Species: Geodia cydonium
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S13179
R;Robitzki, A.; Schroeder, H.C.; Ugarkovic, D.; Kuchino, Y.; Kurelec, B.; Gamulin, V.; Mu
Eur. J. Biochem. 192, 499-506, 1990
A;Title: Regulated expression and phosphorylation of the 23-26-kDa ras protein in the spo
A;Reference number: S13179; MUID:91006138; PMID:2209606
A;Accession: S13179
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-209 <ROB>
A;Note: based on the evidence for Gln-tRNA, the authors translated the codon TAG as Gln;
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;140-143/Region: GTP-binding NKXD motif
F;168-170/Region: GTP-binding SAK/L motif
F;16,17,58,140,141,143,168/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

  Query Match             61.1%;  Score 44;  DB 2;  Length 209;
  Best Local Similarity   53.3%;  Pred. No. 18;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|| ::
Db         10 GGGLVGKSALTLQLV 24


RESULT 6
S40209
tubulin gamma chain - fungus (Cochliobolus heterostrophus)
C;Species: Cochliobolus heterostrophus, Bipolaris maydis
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C;Accession: S40209
R;Parkinson, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlin, M.H.
submitted to the EMBL Data Library, August 1993
A;Description: Phylogenetic analyses using the gamma tubulin gene.
A;Reference number: S40209
A;Accession: S40209
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-265 <PAR>
A;Cross-references: EMBL:X74455; NID:g437988; PIDN:CAA52464.1; PID:g437989
C;Genetics:
A;Introns: 136/3
C;Superfamily: tubulin

  Query Match             61.1%;  Score 44;  DB 2;  Length 265;
  Best Local Similarity   43.8%;  Pred. No. 24;
  Matches    7;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              | : |::|: ||::||
Db        142 GALTRIAADRLHVMTP 157


RESULT 7
S76960
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 27-Oct-2003
C;Accession: S76960
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76960
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <KAN>
A;Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAA18872.1; PID:g165396
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: glutathione S-transferase
```

```
  Query Match             61.1%;  Score 44;  DB 2;  Length 319;
  Best Local Similarity   47.1%;  Pred. No. 29;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || :||: ::  |:|||
Db         20 GGRFVRHDSQFRHWITP 36


RESULT 8
TVDORS
transforming protein ras - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 19-Jan-2001
C;Accession: A01371
R;Reymond, C.D.; Gomer, R.H.; Mehdy, M.C.; Firtel, R.A.
Cell 39, 141-148, 1984
A;Title: Developmental regulation of a Dictyostelium gene encoding a protein homologous
A;Reference number: A01371; MUID:85024887; PMID:6091907
A;Accession: A01371
A;Molecule type: DNA
A;Residues: 1-186 <REY>
A;Cross-references: GB:K02114; NID:g167864; PIDN:AAA33243.1; PID:g167865
C;Genetics:
A;Gene: ras
A;Introns: 25/3; 30/1; 47/1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane protein; methylated carboxyl end; nucleot
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;145-147/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,145/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F;183/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;183/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

  Query Match             59.7%;  Score 43;  DB 1;  Length 186;
  Best Local Similarity   53.3%;  Pred. No. 23;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 9
TVDORA
transforming protein rasG - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 19-Jan-2001
C;Accession: A31456; S21090; S22129
R;Robbins, S.M.; Williams, J.G.; Jermyn, K.A.; Spiegelman, G.B.; Weeks, G.
Proc. Natl. Acad. Sci. U.S.A. 86, 938-942, 1989
A;Title: Growing and developing Dictyostelium cells express different ras genes.
A;Reference number: A31456; MUID:89128893; PMID:2644652
A;Accession: A31456
A;Molecule type: mRNA
A;Residues: 1-189 <ROB1>
A;Cross-references: GB:J04160; NID:g167866; PIDN:AAA33244.1; PID:g167867
R;Robbins, S.M.; Williams, J.G.; Spiegelman, G.B.; Weeks, G.
Biochim. Biophys. Acta 1130, 85-89, 1992
A;Title: Cloning and characterization of the Dictyostelium discoideum rasG genomic seque
A;Reference number: S21090; MUID:92182019; PMID:1339294
A;Accession: S21090
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-189 <ROB2>
A;Cross-references: EMBL:Z11533; NID:g7342; PIDN:CAA77632.1; PID:g7343
C;Genetics:
A;Gene: rasG
A;Introns: 25/3
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane protein; methylated carboxyl end; nucleoti
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat
F;186/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;186/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

  Query Match             59.7%;  Score 43;  DB 1;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 10
S33796
ras protein homolog - slime mold (Physarum polycephalum)
C;Species: Physarum polycephalum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S33796
R;Kozlowski, P.; Fronk, J.; Toczko, K.
Biochim. Biophys. Acta 1173, 357-359, 1993
A;Title: Identification of a ras gene in the slime mold Physarum polycephalum.
A;Reference number: S33796; MUID:93305735; PMID:8318547
A;Accession: S33796
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-189 <KOZ>
A;Cross-references: GB:L10344; GB:Z21495; NID:g1478117; PIDN:AAB05646.1; PID:g310554
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

  Query Match             59.7%;  Score 43;  DB 2;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 11
JC6328
Ras2 protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 21-May-1998 #sequence_revision 29-May-1998 #text_change 19-Jan-2001
C;Accession: JC6328
R;van Es, S.; Kooistra, R.A.; Schaap, P.
Gene 187, 93-97, 1997
A;Title: Two ras genes in Dictyostelium minutum show high sequence homology, but differer
A;Reference number: JC6304; MUID:97225801; PMID:9073071
A;Accession: JC6328
A;Molecule type: DNA
A;Residues: 1-191 <VAN>
C;Comment: This protein is expressed during the entire course of development and is not r
C;Genetics:
A;Gene: ras2
A;Introns: 25/2; 30/1; 65/2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17,57-62,115-118,144-148/Domain: GTP-binding #status predicted <GTB>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
```

```
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

  Query Match             59.7%;  Score 43;  DB 2;  Length 191;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 12
S58220
transforming protein ras-2 - Dictyostelium minutum
C;Species: Dictyostelium minutum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S58220
R;van Es, S.; Kooistra, R.A.; Schaap, P.
submitted to the EMBL Data Library, July 1995
A;Description: Two ras genes in Dictyostelium minutum show high sequence homology, but d
A;Reference number: S58220
A;Accession: S58220
A;Molecule type: DNA
A;Residues: 1-191 <VAN>
A;Cross-references: EMBL:X89037; NID:g929568; PIDN:CAA61434.1; PID:g929569
A;Experimental source: strain 71-2
C;Genetics:
A;Gene: ras2
A;Introns: 25/2; 30/1; 65/2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane protein; methylated carboxyl end; nucleot
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F;188/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;188/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

  Query Match             59.7%;  Score 43;  DB 2;  Length 191;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 13
S55022
transforming protein ras2 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 19-Jan-2001
C;Accession: S55022; S12083
R;Harrison, S.D.; Solomon, N.; Rubin, G.M.
Genetics 139, 1701-1709, 1995
A;Title: A genetic analysis of the 63E-64A genomic region of Drosophila melanogaster: id
A;Reference number: S55020; MUID:95309683; PMID:7789770
A;Accession: S55022
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-192 <HAR>
A;Cross-references: EMBL:U15967; NID:g639707; PIDN:AAB60243.1; PID:g639710
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R;Cohen, N.; Salzberg, A.; Lev, Z.
Oncogene 3, 137-142, 1988
A;Title: A bidirectional promoter is regulating the Drosophila ras2 gene.
A;Reference number: S12083; MUID:88319648; PMID:3412773
A;Accession: S12083
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-27,'VS' <COH>
A;Cross-references: EMBL:X07255; NID:g8402; PIDN:CAA30242.1; PID:g8403
C;Genetics:
A;Gene: ras2
A;Cross-references: FlyBase:FBgn0003206
A;Introns: 27/3; 57/1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop; transforming protein
F;6-121/Domain: translation elongation factor Tu homology <ETU>
F;12-19/Region: nucleotide-binding motif A (P-loop)
F;118-121/Region: GTP-binding NKXD motif
F;148-150/Region: GTP-binding SAK/L motif
F;18,19,37,118,119,121,148/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

  Query Match             59.7%;  Score 43;  DB 2;  Length 192;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSAITIQFI 26


RESULT 14
S32042
GTP-binding protein ras2 - Hydra magnipapillata
C;Species: Hydra magnipapillata
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 19-Jan-2001
C;Accession: JC4573; S32042
R;Bosch, T.C.G.; Benitez, E.; Gellner, K.; Praetzel, G.; Salgado, L.M.
Gene 167, 191-195, 1995
A;Title: Cloning of a ras-related gene from Hydra which responds to head-specific signals
A;Reference number: JC4573; MUID:96144273; PMID:8566776
A;Accession: JC4573
A;Molecule type: mRNA
A;Residues: 1-192 <BOS>
A;Cross-references: EMBL:X70839; NID:g11139; PIDN:CAA50187.1; PID:g11140
A;Experimental source: epithelial cell
C;Comment: This protein is a member of ras protein family, and a key component in recepto
. This protein is highly sensitive to head-specific signals and plays a critical role in
C;Genetics:
A;Gene: ras2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; methylated carboxyl end; nucleotide binding; P-loop
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;37-45/Region: effector
F;58-63/Region: nucleotide-binding motif B
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat
F;189/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;189/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

  Query Match             59.7%;  Score 43;  DB 2;  Length 192;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         15 GGGGVGKSALTIQFI 29


RESULT 15
S38362
Ppras2 protein - slime mold (Physarum polycephalum)
C;Species: Physarum polycephalum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S38362
R;Kozlowski, P.; Tymowska, Z.; Toczko, K.
Biochim. Biophys. Acta 1174, 299-302, 1993
A;Title: Nucleotide and predicted amino acid sequence of a new member of the ras gene fan
A;Reference number: S38362; MUID:93385161; PMID:8373809
A;Accession: S38362
```

```
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-193 <KOZ>
A;Cross-references: GB:L14275; NID:g404808; PIDN:AAC37179.1; PID:g404809
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;6-121/Domain: translation elongation factor Tu homology <ETU>
F;12-19/Region: nucleotide-binding motif A (P-loop)
F;118-121/Region: GTP-binding NKXD motif
F;148-150/Region: GTP-binding SAK/L motif
F;18,19,37,118,119,121,148/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

  Query Match             59.7%;  Score 43;  DB 2;  Length 193;
  Best Local Similarity   53.3%;  Pred. No. 24;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSALTIQLI 26


Search completed: June  2, 2004, 18:13:09
Job time : 4.6124 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:06:18 ; Search time 3.16279 Seconds
(without alignments)
279.877 Million cell updates/sec

Title: US-10-092-367-138
Perfect score: 72
Sequence: 1 GGGXVRXSAXTLHXITP 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	62.5	342	1	RTCA_ECO57	P58127 escherichia
2	45	62.5	1982	1	CHDM_DROME	O97159 drosophila
3	44	61.1	187	1	DEF_CHLTE	Q8kcg7 chlorobium
4	44	61.1	209	1	RAS_GEOCY	P24498 geodia cydo
5	44	61.1	265	1	TBG_COCHE	P40633 cochliobolu
6	44	61.1	338	1	RTCA_ECOLI	P46849 escherichia
7	44	61.1	339	1	RTCA_SALTY	Q8zli0 salmonella
8	43	59.7	187	1	RASD_DICDI	P03967 dictyosteli
9	43	59.7	189	1	RAS1_PHYPO	P34729 physarum po
10	43	59.7	189	1	RASG_DICDI	P15064 dictyosteli
11	43	59.7	192	1	RAS2_DROME	P04388 drosophila
12	43	59.7	192	1	RAS2_HYDMA	P38976 hydra magni
13	43	59.7	193	1	RAS2_PHYPO	P34726 physarum po
14	43	59.7	197	1	RASB_DICDI	P32252 dictyosteli
15	43	59.7	203	1	RAS1_RHIRA	P22278 rhizomucor
16	43	59.7	204	1	RRA2_HUMAN	P17082 homo sapien
17	43	59.7	205	1	RAS3_RHIRA	P22280 rhizomucor
18	43	59.7	215	1	RASL_COPCI	Q05058 coprinus ci
19	43	59.7	216	1	RAS_CRYNE	O74650 cryptococcu
20	43	59.7	217	1	RAS_LENED	P28775 lentinula e
21	43	59.7	218	1	RRAS_HUMAN	P10301 homo sapien
22	43	59.7	218	1	RRAS_MOUSE	P10833 mus musculu
23	43	59.7	290	1	RAS1_CANAL	Q9uqx7 candida alb
24	43	59.7	309	1	RAS1_YEAST	P01119 saccharomyc
25	43	59.7	338	1	RTCA_ECOL6	Q8fcs8 escherichia
26	43	59.7	347	1	RTCA_RALSO	Q8y2v6 ralstonia s
27	42.5	59.0	1723	1	KA93_HUMAN	Q9upq9 homo sapien
28	42	58.3	337	1	RTCA_SULSO	Q97w04 sulfolobus
29	42	58.3	1293	1	SNGP_RAT	Q9quh6 rattus norv
30	42	58.3	1328	1	SNGP_HUMAN	Q96pv0 homo sapien
31	41	56.9	339	1	SRR_MOUSE	Q9qzx7 mus musculu
32	41	56.9	342	1	RTCA_PYRFU	Q8u0n7 pyrococcus
33	41	56.9	399	1	DXR_BORBR	Q7wj88 bordetella
34	41	56.9	399	1	DXR_BORPA	Q7wa54 bordetella
35	41	56.9	399	1	DXR_BORPE	Q7vyc4 bordetella
36	41	56.9	411	1	YB09_METTH	O27181 methanobact
37	41	56.9	460	1	NIFN_RHILO	Q98ap3 rhizobium l
38	41	56.9	531	1	TRPE_ARTGO	P96556 arthrobacte
39	41	56.9	654	1	Z133_HUMAN	P52736 homo sapien
40	41	56.9	715	1	BBS2_BRARE	Q98sp7 brachydanio
41	41	56.9	751	1	Z337_HUMAN	Q9y3m9 homo sapien
42	40	55.6	151	1	BTF3_SCHPO	Q92371 schizosacch
43	40	55.6	183	1	RAP2_HUMAN	P10114 homo sapien
44	40	55.6	183	1	RAP3_HUMAN	P17964 homo sapien
45	40	55.6	210	1	C1IB_HUMAN	Q96iu4 homo sapien

ALIGNMENTS

```
RESULT 1
RTCA_ECO57
ID   RTCA_ECO57     STANDARD;      PRT;   342 AA.
AC   P58127;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-phosphate
DE   cyclase) (RNA cyclase).
GN   RTCA OR Z4778 OR ECS4263.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE005564; AAG58524.1; -.
DR   EMBL; AP002565; BAB37686.1; ALT_INIT.
DR   HSSP; P46849; 1QMH.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Ligase; Complete proteome.
FT   ACT_SITE    308    308       BY SIMILARITY.
FT   DISULFID    307    307       INTERCHAIN (BY SIMILARITY).
SQ   SEQUENCE   342 AA;  36332 MW;  783FE7FAD7160846 CRC64;

  Query Match             62.5%;  Score 45;  DB 1;  Length 342;
  Best Local Similarity   50.0%;  Pred. No. 6.9;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :||::| :||
Db         14 GGGQIMRSALSLSMIT 29


RESULT 2
CHDM_DROME
ID   CHDM_DROME     STANDARD;      PRT;  1982 AA.
AC   O97159; Q8MZ43; Q9VW50;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Chromodomain helicase-DNA-binding protein Mi-2 homolog (dMi-2).
GN   MI-2 OR CG8103.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, AND MUTAGENESIS OF GLY-737.
RX   MEDLINE=99055400; PubMed=9836641;
RA   Kehle J., Beuchle D., Treuheit S., Christen B., Kennison J.A.,
RA   Bienz M., Muller J.;
RT   "dMi-2, a hunchback-interacting protein that functions in Polycomb
RT   repression.";
RL   Science 282:1897-1900(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   REVISIONS.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [4]
RP   SEQUENCE OF 1191-1982 FROM N.A.
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC   -!- FUNCTION: Vital role in development. Protein binds to a portion of
CC       Hunchback (HB) protein that is critical for repression of bithorax
CC       complex (BXC) genes. May also function in polycomb group (PcG)
CC       repression of Hox genes.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
CC   -!- SIMILARITY: Contains 2 chromo domains.
CC   -!- SIMILARITY: Contains 2 PHD-type zinc fingers.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF119716; AAD17276.1; -.
DR   EMBL; AE003515; AAF49099.2; -.
DR   EMBL; AY113368; AAM29373.1; ALT_INIT.
DR   FlyBase; FBgn0013591; Mi-2.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005700; C:polytene chromosome; IDA.
DR   GO; GO:0003682; F:chromatin binding; ISS.
DR   GO; GO:0004386; F:helicase activity; ISS.
DR   GO; GO:0016564; F:transcriptional repressor activity; IGI.
DR   GO; GO:0006333; P:chromatin assembly/disassembly; ISS.
DR   InterPro; IPR000953; Chromo.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002464; DEAH_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000330; SNF2_N.
DR   InterPro; IPR001965; Znf_PHD.
```

```
DR   Pfam; PF00385; chromo; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00628; PHD; 2.
DR   Pfam; PF00176; SNF2_N; 1.
DR   SMART; SM00298; CHROMO; 2.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00249; PHD; 2.
DR   PROSITE; PS00598; CHROMO_1; FALSE_NEG.
DR   PROSITE; PS50013; CHROMO_2; 2.
DR   PROSITE; PS00690; DEAH_ATP_HELICASE; 1.
DR   PROSITE; PS01359; ZF_PHD_1; 2.
DR   PROSITE; PS50016; ZF_PHD_2; 2.
KW   DNA-binding; ATP-binding; Helicase; Nuclear protein; Repeat;
KW   Transcription regulation; Repressor; Zinc-finger.
FT   ZN_FING      377    424       PHD-TYPE 1.
FT   ZN_FING      437    484       PHD-TYPE 2.
FT   DOMAIN       488    566       CHROMO 1.
FT   DOMAIN       612    673       CHROMO 2.
FT   NP_BIND      755    762       ATP (POTENTIAL).
FT   SITE         875    878       DEAH BOX.
FT   DOMAIN        13     16       POLY-GLU.
FT   DOMAIN        70     76       POLY-LYS.
FT   DOMAIN       239    248       POLY-GLU.
FT   DOMAIN      1279   1287       POLY-GLU.
FT   DOMAIN      1672   1677       POLY-ASP.
FT   MUTAGEN      737    737       G->D: IN ALLELE MI-2-5; LARVAL LETHAL.
FT   CONFLICT     101    101       G -> A (IN REF. 1).
SQ   SEQUENCE   1982 AA;  224199 MW;  ED8E256D1AD0AC2F CRC64;

  Query Match             62.5%;  Score 45;  DB 1;  Length 1982;
  Best Local Similarity   50.0%;  Pred. No. 51;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              ||:| :||:| :::||
Db       1541 GGNVDKSATTSNSVTP 1556


RESULT 3
DEF_CHLTE
ID   DEF_CHLTE      STANDARD;      PRT;   187 AA.
AC   Q8KCG7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN   DEF OR CT1454.
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC   -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC       newly synthesized proteins. Requires at least a dipeptide for an
CC       efficient rate of reaction. N-terminal L-methionine is a
CC       prerequisite for activity but the enzyme has broad specificity at
CC       other positions (By similarity).
CC   -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC       methionyl peptide.
CC   -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC   -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE012902; AAM72682.1; -.
DR   TIGR; CT1454; -.
DR   HAMAP; MF_00163; -; 1.
DR   InterPro; IPR000181; Pep_deformylase.
DR   Pfam; PF01327; Pep_deformylase; 1.
DR   PRINTS; PR01576; PDEFORMYLASE.
DR   ProDom; PD003844; Pep_deformylase; 1.
DR   TIGRFAMs; TIGR00079; pept_deformyl; 1.
KW   Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT   ACT_SITE    137    137       BY SIMILARITY.
FT   METAL        94     94       IRON (BY SIMILARITY).
FT   METAL       136    136       IRON (BY SIMILARITY).
FT   METAL       140    140       IRON (BY SIMILARITY).
SQ   SEQUENCE   187 AA;  20909 MW;  1E16EA5077AFC296 CRC64;

  Query Match             61.1%;  Score 44;  DB 1;  Length 187;
  Best Local Similarity   61.5%;  Pred. No. 5.2;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHX 14
              | :||:||:|||:
Db        102 GNVVRPSAITLHY 114


RESULT 4
RAS_GEOCY
ID   RAS_GEOCY      STANDARD;      PRT;   209 AA.
AC   P24498;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein.
OS   Geodia cydonium (Sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Astrophorida; Geodiidae; Geodia.
OX   NCBI_TaxID=6047;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91006138; PubMed=2209606;
RA   Robitzki A., Schroeder H.C., Ugarkovic D., Kuchino Y., Kurelec B.,
RA   Gamulin V., Mueller W.E.G.;
RT   "Regulated expression and phosphorylation of the 23-26-kDa ras
RT   protein in the sponge Geodia cydonium.";
RL   Eur. J. Biochem. 192:499-506(1990).
CC   -!- FUNCTION: This protein is activated by the insulin/insulin
CC       (insulin-like)-receptor system. This transition enables the ras
CC       protein to interact with the lectin-receptor/lectin complex, a
CC       process which ultimately lead to an initiation of an intra-
CC       cellular signal-transduction chain.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- PTM: PHOSPHORYLATED IN THE PRESENCE OF INSULIN.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; M30929; -; NOT_ANNOTATED_CDS.
DR   PIR; S13179; S13179.
DR   HSSP; P01112; 1PLJ.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
KW   GTP-binding; Prenylation; Lipoprotein; Phosphorylation.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      79     83       GTP (BY SIMILARITY).
FT   NP_BIND     140    143       GTP (BY SIMILARITY).
FT   DOMAIN       55     63       EFFECTOR REGION (BY SIMILARITY).
FT   MOD_RES      58     58       PHOSPHORYLATION (POTENTIAL).
FT   LIPID       206    206       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   209 AA;  23854 MW;  C544C43102C8323D CRC64;

  Query Match             61.1%;  Score 44;  DB 1;  Length 209;
  Best Local Similarity   53.3%;  Pred. No. 5.9;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|| ::
Db         10 GGGLVGKSALTLQLV 24


RESULT 5
TBG_COCHE
ID   TBG_COCHE      STANDARD;      PRT;   265 AA.
AC   P40633;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tubulin gamma chain (Gamma tubulin) (Fragment).
OS   Cochliobolus heterostrophus (Drechslera maydis).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC   Pleosporales; Pleosporaceae; Cochliobolus.
OX   NCBI_TaxID=5016;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C5;
RA   Parkinson C., Luo H., Knight A., Ahlquist J., Perlin M.H.;
RL   Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Tubulin is the major constituent of microtubules. Gamma
CC       tubulin is found at microtubule organizing centers (MTOC) such as
CC       the spindle poles or the centrosome, suggesting that it is
CC       involved in the minus-end nucleation of microtubule assembly.
CC   -!- SIMILARITY: Belongs to the tubulin family.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X74455; CAA52464.1; -.
DR   PIR; S40209; S40209.
DR   InterPro; IPR008280; Tub_FtsZ_C.
DR   InterPro; IPR000217; Tubulin.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   Pfam; PF03953; tubulin_C; 1.
DR   PRINTS; PR01161; TUBULIN.
DR   PROSITE; PS00227; TUBULIN; 1.
KW   Microtubule; GTP-binding.
FT   NON_TER       1      1
FT   NP_BIND      77     83       GTP (POTENTIAL).
FT   NON_TER     265    265
SQ   SEQUENCE   265 AA;  29567 MW;  A5DA0C23E7D62DC6 CRC64;

  Query Match             61.1%;  Score 44;  DB 1;  Length 265;
  Best Local Similarity   43.8%;  Pred. No. 7.7;
  Matches    7;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              | : |::|: ||::||
Db        142 GALTRIAADRLHVMTP 157


RESULT 6
RTCA_ECOLI
ID   RTCA_ECOLI     STANDARD;      PRT;   338 AA.
AC   P46849; P46848; Q47349;
DT   01-NOV-1995 (Rel. 32, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-phosphate
DE   cyclase) (RNA cyclase).
GN   RTCA OR B3419/B3420.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE OF 149-339 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=86275993; PubMed=3015733;
RA   Cole S.T., Raibaud O.;
RT   "The nucleotide sequence of the malT gene encoding the positive
RT   regulator of the Escherichia coli maltose regulon.";
RL   Gene 42:201-208(1986).
RN   [3]
RP   REVISION, AND CHARACTERIZATION.
RX   MEDLINE=97327572; PubMed=9184239;
RA   Genschik P., Billy E., Swianiewicz M., Filipowicz W.;
RT   "The human RNA 3'-terminal phosphate cyclase is a member of a new
RT   family of proteins conserved in Eucarya, Bacteria and Archaea.";
RL   EMBO J. 16:2955-2967(1997).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=98411361; PubMed=9738023;
RA   Genschik P., Drabikowski K., Filipowicz W.;
RT   "Characterization of the Escherichia coli RNA 3'-terminal phosphate
RT   cyclase and its sigma54-regulated operon.";
RL   J. Biol. Chem. 273:25516-25526(1998).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RC   STRAIN=K12;
RX   MEDLINE=20139688; PubMed=10673421;
RA   Palm G.J., Billy E., Filipowicz W., Wlodawer A.;
RT   "Crystal structure of RNA 3'-terminal phosphate cyclase, a ubiquitous
RT   enzyme with unusual topology.";
RL   Structure 8:13-23(2000).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing.
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
```

```
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 122 that produces two separate ORFs.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18997; AAA58218.1; ALT_FRAME.
DR   EMBL; U18997; AAA58217.1; ALT_FRAME.
DR   EMBL; AE000418; AAC76445.1; ALT_FRAME.
DR   EMBL; AE000418; AAC76444.1; ALT_FRAME.
DR   EMBL; M13585; AAA83889.1; -.
DR   PDB; 1QMH; 11-JAN-00.
DR   PDB; 1QMI; 11-JAN-00.
DR   EcoGene; EG12938; rtcA.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Ligase; 3D-structure; Complete proteome.
FT   ACT_SITE    308    308       PROBABLE.
FT   DISULFID    307    307       INTERCHAIN.
FT   STRAND        5      8
FT   TURN          9     10
FT   TURN         12     13
FT   HELIX        16     29
FT   STRAND       33     36
FT   TURN         38     41
FT   STRAND       42     42
FT   HELIX        49     62
FT   TURN         63     63
FT   STRAND       65     67
FT   TURN         71     72
FT   STRAND       76     79
FT   STRAND       86     97
FT   HELIX        98    109
FT   TURN        110    111
FT   STRAND      116    123
FT   STRAND      125    126
FT   TURN        127    128
FT   STRAND      129    129
FT   TURN        132    132
FT   HELIX       133    137
FT   TURN        138    138
FT   HELIX       139    145
FT   TURN        146    147
FT   STRAND      149    156
FT   STRAND      159    159
FT   TURN        160    161
FT   STRAND      165    172
FT   STRAND      181    181
FT   STRAND      184    184
FT   STRAND      188    198
FT   HELIX       202    215
FT   STRAND      220    226
FT   HELIX       228    230
FT   STRAND      233    242
FT   STRAND      246    252
FT   TURN        255    256
FT   HELIX       259    275
FT   STRAND      278    278
FT   HELIX       282    295
FT   TURN        296    296
FT   STRAND      299    302
FT   HELIX       307    319
FT   STRAND      325    328
FT   STRAND      333    336
SQ   SEQUENCE   338 AA;  35903 MW;  3450201CB8E40CE7 CRC64;

  Query Match             61.1%;  Score 44;  DB 1;  Length 338;
  Best Local Similarity   50.0%;  Pred. No. 10;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :||::| :||
Db         14 GGGQILRSALSLSMIT 29


RESULT 7
RTCA_SALTY
ID   RTCA_SALTY     STANDARD;      PRT;   339 AA.
AC   Q8ZLI0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-phosphate
DE   cyclase) (RNA cyclase).
GN   RTCA OR STM3518.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008862; AAL22380.1; -.
DR   StyGene; SG?????; rtcA.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; FALSE_NEG.
KW   Ligase; Complete proteome.
FT   ACT_SITE    308    308       BY SIMILARITY.
SQ   SEQUENCE   339 AA;  35457 MW;  182667CD81E31125 CRC64;
```

```
  Query Match             61.1%;  Score 44;  DB 1;  Length 339;
  Best Local Similarity   50.0%;  Pred. No. 10;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :||::| :||
Db         14 GGGQILRSALSLSMIT 29


RESULT 8
RASD_DICDI
ID   RASD_DICDI     STANDARD;      PRT;   187 AA.
AC   P03967;
DT   23-OCT-1986 (Rel. 02, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein rasD (Transforming protein P23).
GN   RASD OR RASA OR RAS.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=85024887; PubMed=6091907;
RA   Reymond C.D., Gomer R.H., Mehdy M.C., Firtel R.A.;
RT   "Developmental regulation of a Dictyostelium gene encoding a protein
RT   homologous to mammalian ras protein.";
RL   Cell 39:141-148(1984).
RN   [2]
RP   REVISIONS.
RC   STRAIN=AX3;
RX   MEDLINE=91115102; PubMed=1703508;
RA   Esch R.K., Firtel R.A.;
RT   "cAMP and cell sorting control the spatial expression of a
RT   developmentally essential cell-type-specific ras gene in
RT   Dictyostelium.";
RL   Genes Dev. 5:9-21(1991).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- DEVELOPMENTAL STAGE: Expressed at a low level in vegetative cells;
CC       not expressed between the onset of development and aggregation,
CC       and is then re-expressed in the multicellular aggregate stages.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02114; AAA33243.1; -.
DR   EMBL; Z11804; CAA77848.1; -.
DR   PIR; A01371; TVDORS.
DR   HSSP; P01112; 1PLK.
DR   DictyBase; DDB0001711; rasD.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      57     61       GTP (BY SIMILARITY).
FT   NP_BIND     116    119       GTP (BY SIMILARITY).
FT   DOMAIN       32     40       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       184    184       S-geranylgeranyl cysteine
FT                                (By similarity).
FT   CONFLICT    137    143       GFNCPFM -> DSLSFH (IN REF. 1).
SQ   SEQUENCE   187 AA;  21202 MW;  7F526253B8316678 CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 187;
  Best Local Similarity   53.3%;  Pred. No. 7.7;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 9
RAS1_PHYPO
ID   RAS1_PHYPO     STANDARD;      PRT;   189 AA.
AC   P34729;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein 1.
GN   RAS1 OR RAS-1.
OS   Physarum polycephalum (Slime mold).
OC   Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
OC   Physarum.
OX   NCBI_TaxID=5791;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LU352;
RX   MEDLINE=93305735; PubMed=8318547;
RA   Kozlowski P., Fronk J., Toczko K.;
RT   "Identification of a ras gene in the slime mold Physarum
RT   polycephalum.";
RL   Biochim. Biophys. Acta 1173:357-359(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M3CVIII;
RX   MEDLINE=96186923; PubMed=8635743;
RA   Trzcinska-Danielewicz J., Kozlowski P., Toczko K.;
RT   "Cloning and genomic sequence of the Physarum polycephalum Pprasl
RT   gene, a homologue of the ras protooncogene.";
RL   Gene 169:143-144(1996).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Inner surface of plasma membrane.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10344; AAB05646.1; -.
DR   EMBL; U10905; AAB06296.1; -.
DR   PIR; S33796; S33796.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      57     61       GTP (BY SIMILARITY).
FT   NP_BIND     116    119       GTP (BY SIMILARITY).
```

```
FT   DOMAIN       32     40       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       186    186       S-geranylgeranyl cysteine
FT                                 (By similarity).
SQ   SEQUENCE   189 AA;  21202 MW;  5EEC8AD372A4CB94 CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 7.8;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 10
RASG_DICDI
ID   RASG_DICDI     STANDARD;      PRT;   189 AA.
AC   P15064;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein rasG.
GN   RASG.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89128893; PubMed=2644652;
RA   Robbins S.M., Williams J.G., Jermyn K.A., Spiegelman G.B., Weeks G.;
RT   "Growing and developing Dictyostelium cells express different ras
RT   genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:938-942(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX2;
RX   MEDLINE=92182019; PubMed=1339294;
RA   Robbins S.M., Williams J.G., Spiegelman G.B., Weeks G.;
RT   "Cloning and characterization of the Dictyostelium discoideum rasG
RT   genomic sequences.";
RL   Biochim. Biophys. Acta 1130:85-89(1992).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04160; AAA33244.1; -.
DR   EMBL; Z11533; CAA77632.1; -.
DR   PIR; A31456; TVDORA.
DR   HSSP; P01112; 1PLK.
DR   DictyBase; DDB0001821; rasG.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      57     61       GTP (BY SIMILARITY).
FT   NP_BIND     116    119       GTP (BY SIMILARITY).
```

```
FT   DOMAIN       32     40       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       186    186       S-geranylgeranyl cysteine
FT                                 (By similarity).
SQ   SEQUENCE   189 AA;  21333 MW;  AFB502319C090899 CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 7.8;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 11
RAS2_DROME
ID   RAS2_DROME     STANDARD;      PRT;   192 AA.
AC   P04388; Q9VZH7;
DT   20-MAR-1987 (Rel. 04, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein 2.
GN   RAS64B OR RAS2.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85187987; PubMed=3921827;
RA   Mozer B., Marlor R., Parkhusrt S., Corces V.G.;
RT   "Characterization and developmental expression of a Drosophila ras
RT   oncogene.";
RL   Mol. Cell. Biol. 5:885-889(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87248071; PubMed=3110012;
RA   Brock H.W.;
RT   "Sequence and genomic structure of ras homologues Dmras85D and
RT   Dmras64B of Drosophila melanogaster.";
RL   Gene 51:129-137(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Iso-1 / Kennison;
RX   MEDLINE=95309683; PubMed=7789770;
RA   Harrison S.D., Solomon N., Rubin G.M.;
RT   "A genetic analysis of the 63E-64A genomic region of Drosophila
RT   melanogaster: identification of mutations in a replication factor C
RT   subunit.";
RL   Genetics 139:1701-1709(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
```

```
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [5]
RP   REVISIONS.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN   [7]
RP   SEQUENCE OF 28-192 FROM N.A.
RX   MEDLINE=84259319; PubMed=6430564;
RA   Neuman-Silberberg F.S., Schejter E., Hoffmann F.M., Shilo B.-Z.;
RT   "The Drosophila ras oncogenes: structure and nucleotide sequence.";
RL   Cell 37:1027-1033(1984).
RN   [8]
RP   SEQUENCE OF 28-192 FROM N.A.
RC   STRAIN=A1;
RX   MEDLINE=20020328; PubMed=10552039;
RA   Gasperini R., Gibson G.;
RT   "Absence of protein polymorphism in the Ras genes of Drosophila
RT   melanogaster.";
RL   J. Mol. Evol. 49:583-590(1999).
RN   [9]
RP   SEQUENCE OF 1-18 AND 44-64 FROM N.A., SPLICE SITES, AND MUTAGENESIS.
RX   MEDLINE=88255843; PubMed=2838380;
RA   Bishop J.G. III, Corces V.G.;
RT   "Expression of an activated ras gene causes developmental
RT   abnormalities in transgenic Drosophila melanogaster.";
RL   Genes Dev. 2:567-577(1988).
RN   [10]
RP   SEQUENCE OF 1-29 FROM N.A.
RX   MEDLINE=88319648; PubMed=3412773;
RA   Cohen N., Salzberg A., Lev Z.;
RT   "A bidirectional promoter is regulating the Drosophila ras2 gene.";
RL   Oncogene 3:137-142(1988).
RN   [11]
RP   CHARACTERIZATION.
RX   MEDLINE=94008534; PubMed=8404533;
RA   Salzberg A., Cohen N., Halachmi N., Kimchie Z., Lev Z.;
RT   "The Drosophila Ras2 and Rop gene pair: a dual homology with a yeast
RT   Ras-like gene and a suppressor of its loss-of-function phenotype.";
RL   Development 117:1309-1319(1993).
CC   -!- FUNCTION: May be involved in endocytic processes and/or other
CC       transport pathways mediated by vesicle trafficking. May interact
CC       functionally with ROP protein. Ras proteins bind GDP/GTP and
CC       possess intrinsic GTPase activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- DEVELOPMENTAL STAGE: A uniform expression is seen in unfertilized
CC       eggs, embryos, larvae, pupae and adult flies. Expression during
CC       embryogenesis is restricted to the CNS and the Garland cells, a
CC       small group of nephrocytes that takes up waste materials from the
CC       hemolymph by endocytosis. In post-embryonic stages, expression is
CC       seen in the larval salivary glands and the CNS, and in the adult
CC       CNS and reproductive systems.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10804; AAA99202.1; ALT_SEQ.
DR   EMBL; M10759; AAA99202.1; JOINED.
DR   EMBL; M10803; AAA99202.1; JOINED.
DR   EMBL; M16431; AAA28849.1; -.
DR   EMBL; M16124; AAA28849.1; JOINED.
DR   EMBL; M16430; AAA28849.1; JOINED.
DR   EMBL; U15967; AAB60243.1; -.
DR   EMBL; AE003480; AAF47845.2; -.
DR   EMBL; AY119135; AAM50995.1; -.
DR   EMBL; K01962; AAA28848.1; ALT_SEQ.
DR   EMBL; K01961; AAA28848.1; JOINED.
DR   EMBL; AF186651; AAF15517.1; -.
DR   EMBL; X12559; CAA31072.1; -.
DR   EMBL; X12558; CAA31071.1; ALT_INIT.
DR   EMBL; X07255; CAA30242.1; -.
DR   PIR; S55022; S55022.
DR   HSSP; P01112; 1PLK.
DR   FlyBase; FBgn0003206; Ras64B.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      12     19       GTP (BY SIMILARITY).
FT   NP_BIND      59     63       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       34     42       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       189    189       S-farnesyl cysteine (By similarity).
FT   MUTAGEN      14     14       G->V: CAUSE DEVELOPMENTAL ABNORMALITIES.
FT   CONFLICT     28     29       SY -> VS (IN REF. 10).
SQ   SEQUENCE   192 AA;  22235 MW;  3F58A3A33E8FDEBC CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 192;
  Best Local Similarity   53.3%;  Pred. No. 8;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSAITIQFI 26


RESULT 12
```

```
RAS2_HYDMA
ID   RAS2_HYDMA     STANDARD;      PRT;   192 AA.
AC   P38976;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein RAS2.
GN   RAS2.
OS   Hydra magnipapillata (Hydra).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6085;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=105;
RX   MEDLINE=96144273; PubMed=8566776;
RA   Bosch T.C.G., Benitez E., Gellner K., Praetzel G., Salgado L.M.;
RT   "Cloning of a ras-related gene from Hydra which responds to head-
RT   specific signals.";
RL   Gene 167:191-195(1995).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- DEVELOPMENTAL STAGE: Ras2 level drops significantly just after the
CC       head is cut. The expression goes up again after 4 to 8 hours.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70839; CAA50187.1; -.
DR   PIR; JC4573; S32042.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      62     66       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       189    189       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   192 AA;  21787 MW;  2DC2ECC18F10C709 CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 192;
  Best Local Similarity   53.3%;  Pred. No. 8;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         15 GGGGVGKSALTIQFI 29


RESULT 13
RAS2_PHYPO
ID   RAS2_PHYPO     STANDARD;      PRT;   193 AA.
AC   P34726;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein 2.
GN   RAS-2.
OS   Physarum polycephalum (Slime mold).
OC   Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
OC   Physarum.
OX   NCBI_TaxID=5791;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LU352;
RX   MEDLINE=93385161; PubMed=8373809;
RA   Kozlowski P., Tymowska Z., Toczko K.;
RT   "Nucleotide and predicted amino acid sequence of a new member of the
RT   ras gene family from the slime mold Physarum polycephalum.";
RL   Biochim. Biophys. Acta 1174:299-302(1993).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Inner surface of plasma membrane.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14275; AAC37179.1; -.
DR   PIR; S38362; S38362.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      12     19       GTP (BY SIMILARITY).
FT   NP_BIND      59     63       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       34     42       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       190    190       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   193 AA;  21634 MW;  4B0B33CD890EE6CD CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 193;
  Best Local Similarity   53.3%;  Pred. No. 8;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSALTIQLI 26


RESULT 14
RASB_DICDI
ID   RASB_DICDI     STANDARD;      PRT;   197 AA.
AC   P32252;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein rasB.
GN   RASB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93205383; PubMed=8455930;
RA   Daniel J.M., Spiegelman G.B., Weeks G.;
RT   "Characterization of a third ras gene, rasB, that is expressed
RT   throughout the growth and development of Dictyostelium discoideum.";
```

```
RL   Oncogene 8:1041-1047(1993).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96622; AAA33246.1; -.
DR   HSSP; P01112; 1PLL.
DR   DictyBase; DDB0001989; rasB.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      13     20       GTP (BY SIMILARITY).
FT   NP_BIND      60     64       GTP (BY SIMILARITY).
FT   NP_BIND     119    122       GTP (BY SIMILARITY).
FT   DOMAIN       35     43       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       194    194       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   197 AA;  22268 MW;  A3D8D3C6846BD9F4 CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 197;
  Best Local Similarity   53.3%;  Pred. No. 8.2;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         13 GGGGVGKSALTIQFI 27


RESULT 15
RAS1_RHIRA
ID   RAS1_RHIRA     STANDARD;      PRT;   203 AA.
AC   P22278;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein 1.
GN   RAS1.
OS   Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Mucor.
OX   NCBI_TaxID=4841;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 1216B;
RX   MEDLINE=91061774; PubMed=1701021;
RA   Casale W.L., McConnell D.G., Wang S.-Y., Lee Y.-J., Linz J.E.;
RT   "Expression of a gene family in the dimorphic fungus Mucor racemosus
RT   which exhibits striking similarity to human ras genes.";
RL   Mol. Cell. Biol. 10:6654-6663(1990).
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- SUBCELLULAR LOCATION: Plasma membrane.
CC   -!- DEVELOPMENTAL STAGE: In all developmental stages analyzed. Its
CC       signal was more intense in sporulating mycelium.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55175; AAA83378.1; -.
DR   PIR; A36365; A36365.
DR   HSSP; P01112; 1PLL.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      17     24       GTP (BY SIMILARITY).
FT   NP_BIND      64     68       GTP (BY SIMILARITY).
FT   NP_BIND     123    126       GTP (BY SIMILARITY).
FT   DOMAIN       39     47       EFFECTOR REGION (PROBABLE).
FT   LIPID       200    200       S-farnesyl cysteine (By similarity).
SQ   SEQUENCE   203 AA;  23236 MW;  52098F53F3966A54 CRC64;

  Query Match             59.7%;  Score 43;  DB 1;  Length 203;
  Best Local Similarity   53.3%;  Pred. No. 8.5;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         17 GGGGVGKSALTIQFI 31


Search completed: June  2, 2004, 18:10:21
Job time : 3.16279 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:09:54 ; Search time 13.1783 Seconds
(without alignments)
407.018 Million cell updates/sec

Title: US-10-092-367-138
Perfect score: 72
Sequence: 1 GGGXVRXSAXTLHXITP 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	66.7	1417	16	Q9HWR8	Q9hwr8 pseudomonas
2	47	65.3	1132	12	Q9WRU1	Q9wru1 macaca mula
3	46	63.9	419	2	Q9RNH3	Q9rnh3 rhodobacter
4	46	63.9	722	5	Q9U0Z4	Q9u0z4 leishmania
5	45	62.5	34	13	Q8QGG0	Q8qgg0 oncorhynchu
6	45	62.5	233	16	Q8XQJ1	Q8xqj1 ralstonia s
7	45	62.5	403	3	O74962	O74962 schizosacch
8	45	62.5	617	16	Q8EJV0	Q8ejv0 shewanella
9	45	62.5	619	10	Q7XUN3	Q7xun3 oryza sativ
10	45	62.5	1043	10	Q7XUB1	Q7xub1 oryza sativ
11	45	62.5	1558	5	Q8IL26	Q8il26 plasmodium
12	45	62.5	2526	5	Q86IS0	Q86is0 dictyosteli
13	44	61.1	202	12	Q919I7	Q919i7 culex nigri
14	44	61.1	319	16	P74752	P74752 synechocyst
15	44	61.1	339	16	Q83MJ7	Q83mj7 shigella fl
16	44	61.1	515	3	Q8NJQ6	Q8njq6 magnaporthe
17	44	61.1	772	5	Q86PA0	Q86pa0 drosophila
18	44	61.1	878	4	Q9GZZ2	Q9gzz2 homo sapien
19	44	61.1	901	4	Q9H195	Q9h195 homo sapien
20	44	61.1	1003	10	Q8H7K0	Q8h7k0 oryza sativ
21	44	61.1	2515	16	Q7UZ67	Q7uz67 rhodopirell
22	43	59.7	113	10	Q7X710	Q7x710 oryza sativ
23	43	59.7	168	5	Q8ITX9	Q8itx9 caenorhabdi
24	43	59.7	176	10	Q9XHV9	Q9xhv9 oryza sativ
25	43	59.7	186	5	Q01208	Q01208 dictyosteli
26	43	59.7	191	5	O97342	O97342 suberites d
27	43	59.7	191	5	Q24471	Q24471 dictyosteli
28	43	59.7	204	11	Q9D0H6	Q9d0h6 mus musculu
29	43	59.7	204	11	Q8C5D1	Q8c5d1 mus musculu
30	43	59.7	210	3	Q9UVQ4	Q9uvq4 cryptococcu
31	43	59.7	210	3	Q9HFU0	Q9hfu0 cryptococcu
32	43	59.7	212	5	O45056	O45056 caenorhabdi
33	43	59.7	213	3	Q9C1I6	Q9c1i6 pisolithus
34	43	59.7	215	3	Q875L4	Q87514 ustilago ma
35	43	59.7	216	3	Q9P8I9	Q9p8i9 suillus bov
36	43	59.7	217	16	Q7U203	Q7u203 mycobacteri
37	43	59.7	218	16	P96280	P96280 mycobacteri
38	43	59.7	275	16	Q9PE73	Q9pe73 xylella fas
39	43	59.7	275	16	Q8PNS1	Q8pns1 xanthomonas
40	43	59.7	275	16	Q8PC46	Q8pc46 xanthomonas
41	43	59.7	275	16	Q87E79	Q87e79 xylella fas
42	43	59.7	289	3	Q9UVU4	Q9uvu4 candida alb
43	43	59.7	453	3	Q7Z9Z2	Q7z9z2 coprinus ci
44	43	59.7	496	13	Q7SX93	Q7sx93 brachydanio
45	43	59.7	574	10	Q8RZD3	Q8rzd3 oryza sativ

ALIGNMENTS

RESULT 1
Q9HWR8
ID Q9HWR8 PRELIMINARY; PRT; 1417 AA.
AC Q9HWR8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable sensor/response regulator hybrid.
GN PA4112.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL; AE004827; AAG07499.1; -.
DR PIR; H83132; H83132.
DR HSSP; P06143; 1AB6.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.

```
DR   GO; GO:0007600; P:sensory perception; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR006189; CHASE.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR003661; His_kinA_N.
DR   InterPro; IPR008207; Hpt.
DR   InterPro; IPR001610; PAC.
DR   InterPro; IPR000700; PAS-assoc_C.
DR   InterPro; IPR000014; PAS_domain.
DR   InterPro; IPR001789; Response_reg.
DR   Pfam; PF03924; CHASE; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00512; HisKA; 1.
DR   Pfam; PF01627; Hpt; 1.
DR   Pfam; PF00785; PAC; 3.
DR   Pfam; PF00989; PAS; 2.
DR   Pfam; PF00072; response_reg; 2.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD000039; Response_reg; 2.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00073; HPT; 1.
DR   SMART; SM00086; PAC; 3.
DR   SMART; SM00091; PAS; 3.
DR   SMART; SM00448; REC; 2.
DR   TIGRFAMs; TIGR00229; sensory_box; 3.
DR   PROSITE; PS50839; CHASE; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50894; HPT; 1.
DR   PROSITE; PS50113; PAC; 3.
DR   PROSITE; PS50112; PAS; 2.
DR   PROSITE; PS50110; RESPONSE_REGULATORY; 2.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase;
KW   Complete proteome.
SQ   SEQUENCE   1417 AA;  153893 MW;  224E2EC9E45EAF2B CRC64;

  Query Match             66.7%;  Score 48;  DB 16;  Length 1417;
  Best Local Similarity   60.0%;  Pred. No. 1.3e+02;
  Matches    9;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              | |:|::||:|||:|
Db       1336 GEGDVQGSAATLHTI 1350


RESULT 2
Q9WRU1
ID   Q9WRU1      PRELIMINARY;      PRT;  1132 AA.
AC   Q9WRU1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SSDNA binding protein (SSDBP).
OS   Macaca mulatta rhadinovirus 17577, and
OS   Macaca mulatta rhadinovirus 26-95.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=83534, 119193;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Macaca mulatta rhadinovirus 17577;
RX   MEDLINE=99174001; PubMed=10074154;
RA   Searles R.P., Bergquam E.P., Axthelm M.K., Wong S.W.;
RT   "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with
RT   similarity to Kaposi's sarcoma-associated herpesvirus/human
RT   herpesvirus 8.";
RL   J. Virol. 73:3040-3053(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Macaca mulatta rhadinovirus 26-95;
RC   STRAIN=MACACA MULATTA RHADINOVIRUS ISOLATE 26-95;
RX   MEDLINE=20173730; PubMed=10708456;
RA   Alexander L., Denekamp L., Knapp A., Auerbach M.R., Damania B.,
RA   Desrosiers R.C.;
RT   "The primary sequence of rhesus monkey rhadinovirus isolate 26-95:
RT   sequence similarities to Kaposi's sarcoma-associated herpesvirus and
RT   rhesus monkey rhadinovirus isolate 17577.";
RL   J. Virol. 74:3388-3398(2000).
DR   EMBL; AF083501; AAD21333.1; -.
DR   EMBL; AF210726; AAF59983.1; -.
DR   GO; GO:0042025; C:host cell nucleus; IEA.
DR   GO; GO:0005743; C:mitochondrial inner membrane; IEA.
DR   GO; GO:0005488; F:binding; IEA.
DR   GO; GO:0003697; F:single-stranded DNA binding; IEA.
DR   GO; GO:0006260; P:DNA replication; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001993; Mitoch_carrier.
DR   InterPro; IPR000635; Viral_DNA_bind.
DR   Pfam; PF00747; viral_DNA_bp; 1.
DR   PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ   SEQUENCE   1132 AA;  126232 MW;  80EF965A16084CDE CRC64;

  Query Match             65.3%;  Score 47;  DB 12;  Length 1132;
  Best Local Similarity   46.7%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GXVRXSAXTLHXITP 17
              |:::: |:|||::||
Db        849 GQIQFYATTLHCLTP 863


RESULT 3
Q9RNH3
ID   Q9RNH3      PRELIMINARY;      PRT;  419 AA.
AC   Q9RNH3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Sensor kinase homolog (Fragment).
GN   CCKA.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B10;
RX   MEDLINE=20105563; PubMed=10639170;
RA   Lang A.S., Beatty J.T.;
RT   "Genetic analysis of a bacterial genetic exchange element: The gene
RT   transfer agent of Rhodobacter capsulatus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:859-864(2000).
CC   -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC       KINASES.
DR   EMBL; AF181079; AAF13178.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0000156; F:two-component response regulator activity; IEA.
DR   GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR   GO; GO:0007600; P:sensory perception; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR003661; His_kinA_N.
DR   InterPro; IPR001789; Response_reg.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00512; HisKA; 1.
DR   Pfam; PF00072; response_reg; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD000039; Response_reg; 1.
```

```
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00448; REC; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   419 AA;  45836 MW;  9A94A5EF348A39FC CRC64;

  Query Match             63.9%;  Score 46;  DB 2;  Length 419;
  Best Local Similarity   46.7%;  Pred. No. 73;
  Matches    7;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||::|:  : ||:|
Db        170 GGGEIRIETENLHLI 184


RESULT 4
Q9U0Z4
ID   Q9U0Z4      PRELIMINARY;      PRT;   722 AA.
AC   Q9U0Z4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   L5883.03.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RA   Murphy L., Harris D., Ivens A.C., Lawson D., Quail M.,
RA   Rajandream M.A., Barrell B.G.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL117384; CAB55614.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   InterPro; IPR000330; SNF2_N.
DR   Pfam; PF00176; SNF2_N; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   722 AA;  74613 MW;  1AFEDBBF764DF361 CRC64;

  Query Match             63.9%;  Score 46;  DB 5;  Length 722;
  Best Local Similarity   46.7%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||: |:||:::|::
Db        227 GGGAPRASANSVHGV 241


RESULT 5
Q8QGG0
ID   Q8QGG0      PRELIMINARY;      PRT;    34 AA.
AC   Q8QGG0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   K-ras (Fragment).
OS   Oncorhynchus gorbuscha (Pink salmon) (Humpback salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8017;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PWS8;
RA   Cronin M.A., Wickliffe J.K., Dunina Y., Baker R.J.;
RT   "K-ras oncogene DNA sequences in pink salmon in streams impacted by
RT   the Exxon Valdez oil spill: no evidence of oil-induced heritable
RT   mutations.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF465435; AAM11562.1; -.
DR   GO; GO:0005525; F:GTP binding; IEA.
DR   GO; GO:0003925; F:small monomeric GTPase activity; IEA.
DR   GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
KW   GTP-binding.
FT   NON_TER      34      34
SQ   SEQUENCE   34 AA;  3702 MW;  41D73D6875AE4F4F CRC64;

  Query Match             62.5%;  Score 45;  DB 13;  Length 34;
  Best Local Similarity   53.3%;  Pred. No. 6.6;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              | |:| :||:|:|:|
Db         10 GAGGVGKSALTIHLI 24


RESULT 6
Q8XQJ1
ID   Q8XQJ1      PRELIMINARY;      PRT;   233 AA.
AC   Q8XQJ1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative transferase protein (EC 2.-.-.-).
GN   RSP1234 OR RS03178.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OG   Plasmid megaplasmid.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646083; CAD18385.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
KW   Transferase; Plasmid; Complete proteome.
SQ   SEQUENCE   233 AA;  24591 MW;  8E11CA0EF79A7291 CRC64;

  Query Match             62.5%;  Score 45;  DB 16;  Length 233;
  Best Local Similarity   53.3%;  Pred. No. 56;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GXVRXSAXTLHXITP 17
              |:  :||:|:|:|||
Db         66 GRKTSSAPTVHLITP 80


RESULT 7
O74962
ID   O74962      PRELIMINARY;      PRT;   403 AA.
```

```
AC   O74962;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative pantothenate kinase.
GN   SPBC4B4.01C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972h-;
RA   Beck A., Reinhardt R, Lyne M., Wood V., Rajandream M.A., Barrell B.G.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL023706; CAA19281.1; -.
DR   PIR; T40473; T40473.
DR   GeneDB_SPombe; SPBC4B4.01c; -.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   InterPro; IPR004567; PanK_eukar.
DR   Pfam; PF03630; Fumble; 1.
DR   TIGRFAMs; TIGR00555; panK_eukar; 1.
KW   Kinase.
SQ   SEQUENCE   403 AA;  44861 MW;  E4574392867BFE20 CRC64;

  Query Match             62.5%;  Score 45;  DB 3;  Length 403;
  Best Local Similarity   37.5%;  Pred. No. 1e+02;
  Matches    6;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              ||  ::|:  :|:|::|
Db        343 GGSFIRNHVQTMHTLT 358


RESULT 8
Q8EJV0
ID   Q8EJV0      PRELIMINARY;      PRT;   617 AA.
AC   Q8EJV0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Metallo-beta-lactamase superfamily protein.
GN   SO0357.
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
RX   MEDLINE=22297686; PubMed=12368813;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA   Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA   Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA   Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA   Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA   Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA   Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT   "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
DR   EMBL; AE015483; AAN53442.1; -.
DR   TIGR; SO0357; -.
DR   GO; GO:0005498; F:sterol carrier activity; IEA.
DR   InterPro; IPR001279; Blactmase-like.
DR   InterPro; IPR003033; SCP2.
DR   Pfam; PF00753; lactamase_B; 1.
KW   Complete proteome.
SQ   SEQUENCE   617 AA;  69623 MW;  5E4C4B25E89FC378 CRC64;

  Query Match             62.5%;  Score 45;  DB 16;  Length 617;
  Best Local Similarity   41.2%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              |||:| :: :|: ::||
Db        490 GGGTVTATPDTIRAMTP 506


RESULT 9
Q7XUN3
ID   Q7XUN3      PRELIMINARY;      PRT;   619 AA.
AC   Q7XUN3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   OSJNBa0064M23.10 protein.
GN   OSJNBA0064M23.10.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA   Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA   Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA   Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA   Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA   Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA   Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA   Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA   Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL606688; CAD41165.1; -.
SQ   SEQUENCE   619 AA;  64918 MW;  4C949277384CF12A CRC64;

  Query Match             62.5%;  Score 45;  DB 10;  Length 619;
  Best Local Similarity   43.8%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHXITP 17
              ||: |: :: ||:|:|
Db        324 GGETRAGSQLLHDISP 339


RESULT 10
Q7XUB1
ID   Q7XUB1      PRELIMINARY;      PRT;  1043 AA.
AC   Q7XUB1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   OSJNBb0032E06.10 protein.
GN   OSJNBB0032E06.10.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA   Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA   Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA   Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA   Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA   Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA   Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA   Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA   Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL663003; CAD41428.1; -.
```

```
SQ   SEQUENCE   1043 AA;  112954 MW;  B2643CAF988180C0 CRC64;

  Query Match             62.5%;  Score 45;  DB 10;  Length 1043;
  Best Local Similarity   47.1%;  Pred. No. 2.9e+02;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              |||:||:|:: | :  |
Db        365 GGGAVRASSRRLEGAVP 381


RESULT 11
Q8IL26
ID   Q8IL26      PRELIMINARY;      PRT;  1558 AA.
AC   Q8IL26;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ser/Thr protein kinase, putative.
GN   PF14_0423.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3D7;
RX   MEDLINE=22255705; PubMed=12368864;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
DR   EMBL; AE014823; AAN37036.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000215; Serpin.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 2.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS00284; SERPIN; 1.
KW   Kinase.
SQ   SEQUENCE   1558 AA;  184250 MW;  6096236F5091014F CRC64;

  Query Match             62.5%;  Score 45;  DB 5;  Length 1558;
  Best Local Similarity   41.2%;  Pred. No. 4.6e+02;
  Matches    7;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              |  :: :||:|:|:|:|
Db        689 GDTQINPSAGTIHYISP 705


RESULT 12
Q86IS0
ID   Q86IS0      PRELIMINARY;      PRT;  2526 AA.
AC   Q86IS0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to Dictyostelium discoideum (Slime mold). ORF DG1040.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RX   MEDLINE=22092622; PubMed=12097910;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA   Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA   Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL   Nature 418:79-85(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RA   Baumgart C.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC116988; AAO51955.1; -.
DR   InterPro; IPR001092; HLH_basic.
DR   InterPro; IPR007196; Not1.
DR   Pfam; PF04054; Not1; 1.
DR   PROSITE; PS00038; HLH_1; 1.
SQ   SEQUENCE   2526 AA;  288862 MW;  2E769231C9F9A183 CRC64;

  Query Match             62.5%;  Score 45;  DB 5;  Length 2526;
  Best Local Similarity   47.1%;  Pred. No. 7.8e+02;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              |||:  :||:  |::||
Db       1110 GGGGNDISADQYHQLTP 1126


RESULT 13
Q919I7
ID   Q919I7      PRELIMINARY;      PRT;  202 AA.
AC   Q919I7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CUN090 putative similar to AcMNPV ORF96.
GN   CUN090.
OS   Culex nigripalpus baculovirus.
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae.
OX   NCBI_TaxID=130556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Florida1997;
RX   MEDLINE=21488685; PubMed=11602755;
RA   Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
RA   Becnel J.J., Rock D.L., Kutish G.F.;
RT   "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus.";
RL   J. Virol. 75:11157-11165(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Florida1997;
RA   Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
RA   Becnel J.J., Rock D.L., Kutish G.F.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF403738; AAK94168.1; -.
DR   InterPro; IPR006883; Baculo_19.
DR   Pfam; PF04798; Baculo_19; 1.
SQ   SEQUENCE   202 AA;  23082 MW;  115F79E4BF667E88 CRC64;

  Query Match             61.1%;  Score 44;  DB 12;  Length 202;
  Best Local Similarity   66.7%;  Pred. No. 70;
  Matches    8;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GGGXVRXSAXTL 12
              |||:||::|:||
Db         30 GGGIVRHAADTL 41


RESULT 14
P74752
ID   P74752      PRELIMINARY;      PRT;   319 AA.
AC   P74752;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein slr0605.
GN   SLR0605.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90917; BAA18872.1; -.
DR   PIR; S76960; S76960.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   319 AA;  36538 MW;  5C4B797C1858EEF1 CRC64;

  Query Match             61.1%;  Score 44;  DB 16;  Length 319;
  Best Local Similarity   47.1%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITP 17
              || :||: ::  |:|||
Db         20 GGRFVRHDSQFRHWITP 36


RESULT 15
Q83MJ7
ID   Q83MJ7      PRELIMINARY;      PRT;   339 AA.
AC   Q83MJ7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chain B, crystal structure of Rna 3'-terminal phosphate cyclase, An
DE   ubiquitous enzyme with Unusual Topology.
GN   SF3442.
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
DR   EMBL; AE015352; AAN44902.1; -.
DR   GO; GO:0003963; F:RNA-3'-phosphate cyclase activity; IEA.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Complete proteome.
SQ   SEQUENCE   339 AA;  36024 MW;  AD7F57EF111CD266 CRC64;

  Query Match             61.1%;  Score 44;  DB 16;  Length 339;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXIT 16
              |||:: :||::| :||
Db         15 GGGQILRSALSLSMIT 30


Search completed: June  2, 2004, 18:12:17
Job time : 15.1783 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 17:58:08 ; Search time 19.1085 Seconds
(without alignments)
251.370 Million cell updates/sec

Title: US-10-092-367-6
Perfect score: 66
Sequence: 1 GGGXVRXSAXTLHXITX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	17	6	ABJ38948	Abj38948 Linear Ga
2	66	100.0	17	6	ABJ38980	Abj38980 Linear Ga
3	66	100.0	17	6	ABJ38850	Abj38850 Linear Ga
4	66	100.0	17	6	ABJ38903	Abj38903 Conopepti
5	60	90.9	95	6	ABJ38902	Abj38902 Conopepti
6	58	87.9	17	6	ABJ38977	Abj38977 Linear Ga
7	56	84.8	17	6	ABJ38976	Abj38976 Linear Ga
8	52	78.8	95	6	ABJ38896	Abj38896 Conopepti
9	50	75.8	17	6	ABJ38897	Abj38897 Conopepti
10	50	75.8	17	6	ABJ38945	Abj38945 Linear Ga
11	50	75.8	17	6	ABJ38847	Abj38847 Linear Ga
12	50	75.8	95	6	ABJ38894	Abj38894 Conopepti
13	48	72.7	17	6	ABJ38846	Abj38846 Linear Ga
14	48	72.7	17	6	ABJ38895	Abj38895 Conopepti
15	48	72.7	17	6	ABJ38944	Abj38944 Linear Ga
16	45	68.2	308	6	ADA32948	Ada32948 Acinetoba
17	45	68.2	339	2	AAW60076	Aaw60076 Escherich
18	44	66.7	791	2	AAW01022	Aaw01022 Multiple
19	44	66.7	805	6	ABP80438	Abp80438 N. gonorr
20	44	66.7	10431	6	ABU54861	Abu54861 Human CA1
21	43	65.2	140	4	ABG18776	Abg18776 Novel hum
22	43	65.2	195	4	ABB57799	Abb57799 Drosophil
23	43	65.2	203	2	AAW82591	Aaw82591 Human TC2
24	43	65.2	203	6	ABR41057	Abr41057 Human MAP
25	43	65.2	204	2	AAR77647	Aar77647 TC21 muta
26	43	65.2	204	6	ABR41056	Abr41056 Human MAP
27	43	65.2	210	2	AAY42695	Aay42695 Human R-R
28	43	65.2	213	3	AAB07940	Aab07940 Amino aci
29	43	65.2	218	5	AAU75736	Aau75736 Human rel
30	43	65.2	218	6	ABU62885	Abu62885 Ras-famil
31	43	65.2	252	4	ABG18778	Abg18778 Novel hum
32	43	65.2	288	4	AAG74576	Aag74576 Human col
33	43	65.2	439	6	ABU00221	Abu00221 Human nov
34	43	65.2	740	4	AAB99359	Aab99359 Human R-R
35	43	65.2	740	5	ABB06727	Abb06727 Human R-R
36	43	65.2	740	5	ABB06737	Abb06737 Human R-R
37	43	65.2	764	6	ABU00312	Abu00312 Human nov
38	42	63.6	23	6	ABO12068	Abo12068 Human zin
39	42	63.6	69	3	AAG27177	Aag27177 Zea mays
40	42	63.6	97	3	AAG27176	Aag27176 Zea mays
41	42	63.6	124	3	AAG27175	Aag27175 Zea mays
42	42	63.6	146	4	AAM25718	Aam25718 Human pro
43	42	63.6	146	6	ABO00934	Abo00934 Polypepti
44	42	63.6	175	4	AAU65823	Aau65823 Propionib
45	42	63.6	175	6	ABM62342	Abm62342 Propionib

ALIGNMENTS

RESULT 1
ABJ38948
ID ABJ38948 standard; peptide; 17 AA.
XX
AC ABJ38948;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt5 SEQ ID No 138.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 4
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 7
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 10
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 14
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;

PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 43; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 66; DB 6; Length 17;
Best Local Similarity 94.1%; Pred. No. 0.00057;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              ||||||||||||||||:
Db          1 GGGXVRXSAXTLHXITP 17
```

RESULT 2
ABJ38980
ID ABJ38980 standard; peptide; 17 AA.
XX
AC ABJ38980;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt5 SEQ ID No 170.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 44; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 66; DB 6; Length 17;
Best Local Similarity 70.6%; Pred. No. 0.00057;
Matches 12; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
```

```
                     |||:||:||:|||:||:
Db          1 GGGEVRESAETLHEITP 17


RESULT 3
ABJ38850
ID   ABJ38850 standard; peptide; 17 AA.
XX
AC   ABJ38850;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide SEQ ID No 6.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   4
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   7
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   10
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   14
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   17
FT                   /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Claim 1; Page 48; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             100.0%;  Score 66;  DB 6;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.00057;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||||||||||||||||
Db          1 GGGXVRXSAXTLHXITX 17


RESULT 4
ABJ38903
ID   ABJ38903 standard; peptide; 17 AA.
XX
AC   ABJ38903;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide toxin peptide Bt5 SEQ ID No 74.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm; toxin.
XX
OS   Conus betulinus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   4
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   7
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   10
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   14
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   17
```

```
FT                   /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Example 7; Page 33; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a toxin sequence of a linear gamma-
CC   carboxyglutamate rich conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             100.0%;  Score 66;  DB 6;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.00057;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||||||||||||||||
Db          1 GGGXVRXSAXTLHXITX 17


RESULT 5
ABJ38902
ID   ABJ38902 standard; protein; 95 AA.
XX
AC   ABJ38902;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide conotoxin protein Bt5 SEQ ID No 73.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
DR   N-PSDB; ABT43476.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Claim 5; Page 33; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
```

```
C   disorder is anxiety, major depression, manic-depressive illness,
C   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
C   bipolar disorder, unipolar depression, dysthymia or seasonal effective
C   disorder. The conotoxin peptides are also useful for controlling
C   nematodes or parasitic worms by applying the peptides to the locus to be
C   protected. This sequence represents a linear gamma-carboxyglutamate rich
C   conotoxin protein of the invention
X
Q   Sequence 95 AA;

 Query Match             90.9%;  Score 60;  DB 6;  Length 95;
 Best Local Similarity   68.8%;  Pred. No. 0.049;
 Matches   11;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

y          2 GGXVRXSAXTLHXITX 17
             ||:||:||:|||:||:
b         80 GGEVRESAETLHEITP 95


ESULT 6
ABJ38977
D   ABJ38977 standard; peptide; 17 AA.
AX
AC   ABJ38977;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide Bt2 SEQ ID No 167.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Example 7; Page 44; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             87.9%;  Score 58;  DB 6;  Length 17;
  Best Local Similarity   64.7%;  Pred. No. 0.012;
  Matches   11;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              || :||:||:|||:||:
Db          1 GGEEVRESAETLHEITP 17


RESULT 7
ABJ38976
ID   ABJ38976 standard; peptide; 17 AA.
XX
AC   ABJ38976;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide Bt1 SEQ ID No 166.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
```

XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 44; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 84.8%; Score 56; DB 6; Length 17;
Best Local Similarity 58.8%; Pred. No. 0.027;
Matches 10; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || :||:||:|||::|:
Db          1 GGEEVRESAETLHELTP 17
```

RESULT 8
ABJ38896
ID ABJ38896 standard; protein; 95 AA.
XX
AC ABJ38896;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Bt2 SEQ ID No 64.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43473.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 32; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 95 AA;

Query Match 78.8%; Score 52; DB 6; Length 95;
Best Local Similarity 62.5%; Pred. No. 1.1;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GGXVRXSAXTLHXITX 17
              | :||:||:|||:||:
Db         80 GEEVRESAETLHEITP 95
```

RESULT 9
ABJ38897
ID ABJ38897 standard; peptide; 17 AA.
XX
AC ABJ38897;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide toxin peptide Bt2 SEQ ID No 65.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm; toxin.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 3
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 4
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 7
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 10
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 14
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 17
FT /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 32; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a toxin sequence of a linear gamma-
CC carboxyglutamate rich conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 75.8%; Score 50; DB 6; Length 17;
Best Local Similarity 94.1%; Pred. No. 0.27;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||||||
Db          1 GGXXVRXSAXTLHXITX 17
```

RESULT 10
ABJ38945
ID ABJ38945 standard; peptide; 17 AA.
XX
AC ABJ38945;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt2 SEQ ID No 135.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 3

```
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   4
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   7
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   10
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   14
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Example 7; Page 43; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             75.8%;  Score 50;  DB 6;  Length 17;
  Best Local Similarity   88.2%;  Pred. No. 0.27;
  Matches   15;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||||:
Db          1 GGXXVRXSAXTLHXITP 17


RESULT 11
ABJ38847
ID   ABJ38847 standard; peptide; 17 AA.
XX
AC   ABJ38847;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Linear Gamma-carboxyglutamate rich conotoxin peptide SEQ ID No 3.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   4
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   7
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   10
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   14
FT                   /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT   Modified-site   17
FT                   /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Claim 1; Page 48; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
```

CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 75.8%; Score 50; DB 6; Length 17;
Best Local Similarity 94.1%; Pred. No. 0.27;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||||||
Db          1 GGXXVRXSAXTLHXITX 17
```

RESULT 12
ABJ38894
ID ABJ38894 standard; protein; 95 AA.
XX
AC ABJ38894;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Bt1 SEQ ID No 61.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43472.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 31; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 95 AA;

Query Match 75.8%; Score 50; DB 6; Length 95;
Best Local Similarity 56.2%; Pred. No. 2.3;
Matches 9; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GGXVRXSAXTLHXITX 17
              | :||:||:|||::|:
Db         80 GEEVRESAETLHELTP 95
```

RESULT 13
ABJ38846
ID ABJ38846 standard; peptide; 17 AA.
XX
AC ABJ38846;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide SEQ ID No 2.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;

KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 3
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 4
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 7
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 10
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 14
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 17
FT /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 1; Page 48; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 72.7%; Score 48; DB 6; Length 17;
Best Local Similarity 88.2%; Pred. No. 0.59;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||:||
Db          1 GGXXVRXSAXTLHXLTX 17
```

RESULT 14
ABJ38895
ID ABJ38895 standard; peptide; 17 AA.
XX
AC ABJ38895;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide toxin peptide Bt1 SEQ ID No 62.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm; toxin.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 3
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 4
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 7
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 10
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 14
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 17
FT /note= "Residue is optionally Pro or hydroxy-Pro"
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX

PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 31; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a toxin sequence of a linear gamma-
CC carboxyglutamate rich conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 72.7%; Score 48; DB 6; Length 17;
Best Local Similarity 88.2%; Pred. No. 0.59;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||:||
Db          1 GGXXVRXSAXTLHXLTX 17
```

RESULT 15
ABJ38944
ID ABJ38944 standard; peptide; 17 AA.
XX
AC ABJ38944;
XX
DT 09-OCT-2003 (first entry)
XX
DE Linear Gamma-carboxyglutamate rich conotoxin peptide Bt1 SEQ ID No 134.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin peptide;
KW parasitic worm.
XX
OS Conus betulinus.
XX
FH Key Location/Qualifiers
FT Modified-site 3
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 4
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 7
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 10
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
FT Modified-site 14
FT /note= "Residue is optionally Glu or gamma-carboxy-Glu"
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Example 7; Page 43; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as

CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin peptide of the invention
XX
SQ Sequence 17 AA;

Query Match 72.7%; Score 48; DB 6; Length 17;
Best Local Similarity 82.4%; Pred. No. 0.59;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||:|:
Db          1 GGXXVRXSAXTLHXLTP 17
```

Search completed: June 2, 2004, 18:09:43
Job time : 21.1085 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:12:24 ; Search time 5.53488 Seconds
(without alignments)
158.565 Million cell updates/sec

Title: US-10-092-367-6
Perfect score: 66
Sequence: 1 GGGXVRXSAXTLHXITX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	72.7	1441	4	US-09-252-991A-28143	Sequence 28143, A
2	45	68.2	308	4	US-09-328-352-4235	Sequence 4235, Ap
3	45	68.2	415	4	US-09-252-991A-31348	Sequence 31348, A
4	44	66.7	206	2	US-08-531-525-50	Sequence 50, Appl
5	44	66.7	206	2	US-08-718-270A-50	Sequence 50, Appl
6	44	66.7	791	1	US-08-394-880B-2	Sequence 2, Appli
7	43	65.2	187	3	US-09-078-317-11	Sequence 11, Appl
8	43	65.2	188	2	US-08-531-525-47	Sequence 47, Appl
9	43	65.2	188	2	US-08-718-270A-47	Sequence 47, Appl
10	43	65.2	204	3	US-09-078-317-14	Sequence 14, Appl
11	43	65.2	204	4	US-09-454-818-14	Sequence 14, Appl
12	43	65.2	210	4	US-09-053-374A-7	Sequence 7, Appli
13	43	65.2	213	4	US-09-503-505A-3	Sequence 3, Appli
14	43	65.2	215	2	US-08-531-525-49	Sequence 49, Appl
15	43	65.2	215	2	US-08-718-270A-49	Sequence 49, Appl
16	42	63.6	350	4	US-09-821-736-2	Sequence 2, Appli
17	42	63.6	402	4	US-09-543-681A-7141	Sequence 7141, Ap
18	41	62.1	179	2	US-08-531-525-38	Sequence 38, Appl
19	41	62.1	179	2	US-08-718-270A-38	Sequence 38, Appl
20	41	62.1	183	2	US-08-531-525-39	Sequence 39, Appl
21	41	62.1	183	2	US-08-718-270A-39	Sequence 39, Appl
22	41	62.1	183	4	US-09-482-520A-8	Sequence 8, Appli
23	41	62.1	438	4	US-09-252-991A-27582	Sequence 27582, A
24	40	60.6	142	4	US-09-621-976-4099	Sequence 4099, Ap
25	40	60.6	160	4	US-09-489-039A-11449	Sequence 11449, A
26	40	60.6	210	4	US-09-247-155-117	Sequence 117, App
27	40	60.6	605	4	US-09-540-236-2056	Sequence 2056, Ap
28	40	60.6	609	4	US-09-198-452A-579	Sequence 579, App
29	40	60.6	757	4	US-09-252-991A-25918	Sequence 25918, A
30	39	59.1	154	4	US-09-540-236-3363	Sequence 3363, Ap
31	39	59.1	190	2	US-08-531-525-45	Sequence 45, Appl
32	39	59.1	190	2	US-08-718-270A-45	Sequence 45, Appl
33	39	59.1	196	4	US-09-252-991A-26430	Sequence 26430, A
34	39	59.1	217	3	US-09-078-317-9	Sequence 9, Appli
35	39	59.1	217	4	US-09-454-818-9	Sequence 9, Appli
36	39	59.1	219	3	US-09-078-317-10	Sequence 10, Appl
37	39	59.1	219	4	US-09-454-818-10	Sequence 10, Appl
38	39	59.1	334	4	US-09-489-039A-10472	Sequence 10472, A
39	39	59.1	422	4	US-09-252-991A-22959	Sequence 22959, A
40	39	59.1	433	4	US-09-252-991A-23792	Sequence 23792, A
41	39	59.1	446	4	US-09-252-991A-23809	Sequence 23809, A
42	39	59.1	478	4	US-09-252-991A-23438	Sequence 23438, A
43	39	59.1	2293	3	US-09-368-590-2	Sequence 2, Appli
44	38.5	58.3	478	4	US-09-328-352-5547	Sequence 5547, Ap
45	38	57.6	45	2	US-08-935-450-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-28143
; Sequence 28143, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28143
; LENGTH: 1441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28143

Query Match 72.7%; Score 48; DB 4; Length 1441;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              | |:|::||:|||:|
Db       1360 GEGDVQGSAATLHTI 1374
```

RESULT 2
US-09-328-352-4235
; Sequence 4235, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4235
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4235

Query Match 68.2%; Score 45; DB 4; Length 308;

```
  Best Local Similarity   35.3%;  Pred. No. 22;
  Matches    6;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: :: : ||::|:
Db        240 GGGIINHTIPLLHHVTE 256


RESULT 3
US-09-252-991A-31348
; Sequence 31348, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 31348
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31348

  Query Match             68.2%;  Score 45;  DB 4;  Length 415;
  Best Local Similarity   53.3%;  Pred. No. 32;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              | |:||:||: ||::
Db        189 GAGPVRASAVLLHPM 203


RESULT 4
US-08-531-525-50
; Sequence 50, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  206 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Geodia cydonium
US-08-531-525-50

  Query Match             66.7%;  Score 44;  DB 2;  Length 206;
  Best Local Similarity   53.3%;  Pred. No. 20;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|| ::
Db          9 GGGLVGKSALTLQLV 23


RESULT 5
US-08-718-270A-50
; Sequence 50, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/718,270A
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/531,525
;      FILING DATE:  21-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,091
;      FILING DATE:  21-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  78-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  206 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
```

```
;    ORIGINAL SOURCE:
;      ORGANISM:  Geodia cydonium
US-08-718-270A-50

  Query Match             66.7%;  Score 44;  DB 2;  Length 206;
  Best Local Similarity   53.3%;  Pred. No. 20;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|| ::
Db          9 GGGLVGKSALTLQLV 23


RESULT 6
US-08-394-880B-2
; Sequence 2, Application US/08394880B
; Patent No. 5705352
;  GENERAL INFORMATION:
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Skatrud, Paul L.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene Of
;    TITLE OF INVENTION:  Aspergillus Fumigatus
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company/Patent Division
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  US
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/394,880B
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant G., Thomas
;      REGISTRATION NUMBER:  35784
;      REFERENCE/DOCKET NUMBER:  X-9682
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-2459
;      TELEFAX:  (317) 277-1917
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  791 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-394-880B-2

  Query Match             66.7%;  Score 44;  DB 1;  Length 791;
  Best Local Similarity   41.2%;  Pred. No. 1e+02;
  Matches    7;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:|:: |:|: ::|:
Db        431 GGGMVQSGAITIGELTS 447


RESULT 7
US-09-078-317-11
; Sequence 11, Application US/09078317
; Patent No. 6017710
;  GENERAL INFORMATION:
;    APPLICANT:  Allen, Maxine J.
;    APPLICANT:  Rutter, Marc
;    APPLICANT:  Buckler, Alan J.
;    TITLE OF INVENTION:  RAQ Genes and Their Uses
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Ave, Suite 200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/078,317
;      FILING DATE:  13-MAY-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Francis, Carol L
;      REGISTRATION NUMBER:  36,513
;      REFERENCE/DOCKET NUMBER:  SEQ-18P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-327-3400
;      TELEFAX:  650-327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  187 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6017710e
US-09-078-317-11

  Query Match             65.2%;  Score 43;  DB 3;  Length 187;
  Best Local Similarity   53.3%;  Pred. No. 25;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 8
US-08-531-525-47
; Sequence 47, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
```

; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
US-08-531-525-47

Query Match 65.2%; Score 43; DB 2; Length 188;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
||||:| :||:|: :|
Db 9 GGGGVGKSALTIQLI 23

RESULT 9
US-08-718-270A-47
; Sequence 47, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
US-08-718-270A-47

Query Match 65.2%; Score 43; DB 2; Length 188;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
||||:| :||:|: :|
Db 9 GGGGVGKSALTIQLI 23

RESULT 10
US-09-078-317-14
; Sequence 14, Application US/09078317
; Patent No. 6017710
; GENERAL INFORMATION:
; APPLICANT: Allen, Maxine J.
; APPLICANT: Rutter, Marc
; APPLICANT: Buckler, Alan J.
; TITLE OF INVENTION: RAQ Genes and Their Uses
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,317
; FILING DATE: 13-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Francis, Carol L
; REGISTRATION NUMBER: 36,513
; REFERENCE/DOCKET NUMBER: SEQ-18P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6017710e
US-09-078-317-14

Query Match 65.2%; Score 43; DB 3; Length 204;
Best Local Similarity 53.3%; Pred. No. 28;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
||||:| :||:|: :|
Db 21 GGGGVGKSALTIQFI 35

```
RESULT 11
US-09-454-818-14
; Sequence 14, Application US/09454818
; Patent No. 6383792
; GENERAL INFORMATION:
;  APPLICANT:  Allen, Maxine J.
;  APPLICANT:  Rutter, Marc
;  APPLICANT:  Buckler, Alan J.
;  TITLE OF INVENTION:  RAQ Genes and Their Uses
;  FILE REFERENCE:  AXYS-018DIV
;  CURRENT APPLICATION NUMBER: US/09/454,818
;  CURRENT FILING DATE:  1999-12-03
;  PRIOR APPLICATION NUMBER:  09/078,317
;  PRIOR FILING DATE:  1998-05-13
;  NUMBER OF SEQ ID NOS:  16
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-454-818-14

  Query Match             65.2%;  Score 43;  DB 4;  Length 204;
  Best Local Similarity   53.3%;  Pred. No. 28;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         21 GGGGVGKSALTIQFI 35


RESULT 12
US-09-053-374A-7
; Sequence 7, Application US/09053374A
; Patent No. 6462177
;  GENERAL INFORMATION:
;    APPLICANT:  YEN, KWANG-MU
;    TITLE OF INVENTION:  MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/053,374A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-514
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-053-374A-7

  Query Match             65.2%;  Score 43;  DB 4;  Length 210;
  Best Local Similarity   53.3%;  Pred. No. 29;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         28 GGGGVGKSALTIQFI 42


RESULT 13
US-09-503-505A-3
; Sequence 3, Application US/09503505A
; Patent No. 6387688
; GENERAL INFORMATION:
;  APPLICANT: SHISHIDO, KAZUO
;  APPLICANT:  KAJIWARA, SUSUMU
;  APPLICANT:  TSUKAMOTO AKIRA
;  TITLE OF INVENTION: DNA FRAGMENTS HAVING BASIDIOMYCETE-DERIVED PROMOTER
;  TITLE OF INVENTION:  ACTIVITY AND EXPRESSION OF FOREIGN GENES UNDER
;  TITLE OF INVENTION:  CONTROL OF THE PROMOTER ACTIVITY
;  FILE REFERENCE: 04853.0039
;  CURRENT APPLICATION NUMBER: US/09/503,505A
;  CURRENT FILING DATE:  2000-02-14
;  PRIOR APPLICATION NUMBER: JP 36367/1999
;  PRIOR FILING DATE: 1999-02-15
;  PRIOR APPLICATION NUMBER: JP 93777/1999
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 3
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Coriolus hirsutus
US-09-503-505A-3

  Query Match             65.2%;  Score 43;  DB 4;  Length 213;
  Best Local Similarity   53.3%;  Pred. No. 30;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         14 GGGGVGKSALTIQFI 28


RESULT 14
US-08-531-525-49
; Sequence 49, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 58406831e, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Coprinus cinereus
US-08-531-525-49

Query Match 65.2%; Score 43; DB 2; Length 215;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         17 GGGGVGKSALTIQFI 31
```

RESULT 15
US-08-718-270A-49
; Sequence 49, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Coprinus cinereus
US-08-718-270A-49

Query Match 65.2%; Score 43; DB 2; Length 215;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         17 GGGGVGKSALTIQFI 31
```

Search completed: June 2, 2004, 18:13:55
Job time : 6.53488 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:13:14 ; Search time 14.3643 Seconds
(without alignments)
332.960 Million cell updates/sec

Title: US-10-092-367-6
Perfect score: 66
Sequence: 1 GGGXVRXSAXTLHXITX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	17	12	US-10-092-367-6	Sequence 6, Appli
2	66	100.0	17	12	US-10-092-367-74	Sequence 74, Appl
3	66	100.0	17	12	US-10-092-367-138	Sequence 138, App
4	66	100.0	17	12	US-10-092-367-170	Sequence 170, App
5	60	90.9	95	12	US-10-092-367-73	Sequence 73, Appl
6	58	87.9	17	12	US-10-092-367-167	Sequence 167, App
7	56	84.8	17	12	US-10-092-367-166	Sequence 166, App
8	52	78.8	95	12	US-10-092-367-64	Sequence 64, Appl
9	50	75.8	17	12	US-10-092-367-3	Sequence 3, Appli
10	50	75.8	17	12	US-10-092-367-65	Sequence 65, Appl
11	50	75.8	17	12	US-10-092-367-135	Sequence 135, App
12	50	75.8	95	12	US-10-092-367-61	Sequence 61, Appl
13	48	72.7	17	12	US-10-092-367-2	Sequence 2, Appli
14	48	72.7	17	12	US-10-092-367-62	Sequence 62, Appl
15	48	72.7	17	12	US-10-092-367-134	Sequence 134, App
16	45	68.2	333	12	US-10-424-599-244796	Sequence 244796,
17	43	65.2	102	12	US-10-424-599-200067	Sequence 200067,
18	43	65.2	192	12	US-10-424-599-282377	Sequence 282377,
19	43	65.2	203	14	US-10-197-666A-84	Sequence 84, Appl
20	43	65.2	204	14	US-10-197-666A-82	Sequence 82, Appl
21	43	65.2	204	16	US-10-408-765A-1241	Sequence 1241, Ap
22	43	65.2	218	10	US-09-873-546-14	Sequence 14, Appl
23	43	65.2	218	13	US-10-067-813-17	Sequence 17, Appl
24	43	65.2	218	16	US-10-408-765A-690	Sequence 690, App
25	43	65.2	231	15	US-10-369-493-5603	Sequence 5603, Ap
26	43	65.2	288	14	US-10-106-698-5350	Sequence 5350, Ap
27	43	65.2	309	9	US-09-801-368-286	Sequence 286, App
28	43	65.2	309	15	US-10-369-493-22369	Sequence 22369, A
29	43	65.2	740	12	US-10-344-404-23	Sequence 23, Appl
30	42	63.6	146	12	US-10-296-115-1233	Sequence 1233, Ap
31	42	63.6	189	15	US-10-094-749-1717	Sequence 1717, Ap
32	42	63.6	377	15	US-10-051-874-147	Sequence 147, App
33	42	63.6	386	12	US-10-243-552-408	Sequence 408, App
34	42	63.6	402	12	US-10-245-752-34	Sequence 34, Appl
35	42	63.6	402	12	US-10-245-859-34	Sequence 34, Appl
36	42	63.6	402	14	US-10-245-103-34	Sequence 34, Appl
37	42	63.6	402	14	US-10-245-107-34	Sequence 34, Appl
38	42	63.6	402	14	US-10-245-143-34	Sequence 34, Appl
39	42	63.6	402	14	US-10-245-771-34	Sequence 34, Appl
40	42	63.6	402	14	US-10-245-851-34	Sequence 34, Appl
41	42	63.6	402	14	US-10-245-883-34	Sequence 34, Appl
42	42	63.6	402	14	US-10-237-535-34	Sequence 34, Appl
43	42	63.6	402	14	US-10-238-183-34	Sequence 34, Appl
44	42	63.6	402	14	US-10-238-283-34	Sequence 34, Appl
45	42	63.6	402	14	US-10-238-370-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-10-092-367-6
; Sequence 6, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Xaa at residues 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; Xaa
; OTHER INFORMATION: at residue 17 is Pro or hydroxy-Pro
US-10-092-367-6

Query Match 100.0%; Score 66; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITX 17
||||||||||||||||||

```
Db          1 GGGXVRXSAXTLHXITX 17


RESULT 2
US-10-092-367-74
; Sequence 74, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; Xaa
;   OTHER INFORMATION: at residue 17 is Pro or hydroxy-Pro
US-10-092-367-74

  Query Match           100.0%;  Score 66;  DB 12;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 0.002;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||||||||||||||||
Db          1 GGGXVRXSAXTLHXITX 17


RESULT 3
US-10-092-367-138
; Sequence 138, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 138
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa is Glu or gamma-carboxy-Glu
US-10-092-367-138

  Query Match           100.0%;  Score 66;  DB 12;  Length 17;
  Best Local Similarity  94.1%;  Pred. No. 0.002;
  Matches   16;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||||||||||||||||:
Db          1 GGGXVRXSAXTLHXITP 17


RESULT 4
US-10-092-367-170
; Sequence 170, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
US-10-092-367-170

  Query Match           100.0%;  Score 66;  DB 12;  Length 17;
  Best Local Similarity  70.6%;  Pred. No. 0.002;
  Matches   12;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:||:||:|||:||:
Db          1 GGGEVRESAETLHEITP 17


RESULT 5
US-10-092-367-73
; Sequence 73, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
;   LENGTH: 95
;   TYPE: PRT
;   ORGANISM: Conus betulinus
US-10-092-367-73
```

Query Match 90.9%; Score 60; DB 12; Length 95;
Best Local Similarity 68.8%; Pred. No. 0.12;
Matches 11; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 GGXVRXSAXTLHXITX 17
              ||:||:||:|||:||:
Db         80 GGEVRESAETLHEITP 95
```

RESULT 6
US-10-092-367-167
; Sequence 167, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 167
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-167

Query Match 87.9%; Score 58; DB 12; Length 17;
Best Local Similarity 64.7%; Pred. No. 0.035;
Matches 11; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || :||:||:|||:||:
Db          1 GGEEVRESAETLHEITP 17
```

RESULT 7
US-10-092-367-166
; Sequence 166, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 166
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-166

Query Match 84.8%; Score 56; DB 12; Length 17;
Best Local Similarity 58.8%; Pred. No. 0.072;
Matches 10; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || :||:||:|||::|:
Db          1 GGEEVRESAETLHELTP 17
```

RESULT 8
US-10-092-367-64
; Sequence 64, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 64
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-64

Query Match 78.8%; Score 52; DB 12; Length 95;
Best Local Similarity 62.5%; Pred. No. 2.2;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GGXVRXSAXTLHXITX 17
              | :||:||:|||:||:
Db         80 GEEVRESAETLHEITP 95
```

RESULT 9
US-10-092-367-3
; Sequence 3, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)

```
;   OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
;   OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-3

  Query Match             75.8%;  Score 50;  DB 12;  Length 17;
  Best Local Similarity   94.1%;  Pred. No. 0.62;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||||||
Db          1 GGXXVRXSAXTLHXITX 17


RESULT 10
US-10-092-367-65
; Sequence 65, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 65
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Conus betulinus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
;   OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
| US-10-092-367-65

  Query Match             75.8%;  Score 50;  DB 12;  Length 17;
  Best Local Similarity   94.1%;  Pred. No. 0.62;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

| Qy          1 GGGXVRXSAXTLHXITX 17
              || |||||||||||||||
| Db          1 GGXXVRXSAXTLHXITX 17


| RESULT 11
| US-10-092-367-135
  ; Sequence 135, Application US/10092367
  ; Publication No. US20030065138A1
  ; GENERAL INFORMATION:
  ;  APPLICANT: University of Utah Research Foundation
  ;  APPLICANT:  Cognetix, Inc.
  ;  APPLICANT:  Olivera, Baldomero M
  ;  APPLICANT:  McIntosh, J. Michael
  ;  APPLICANT:  Garrett, James E.
  ;  APPLICANT:  Walker, Craig S.
  ;  APPLICANT:  Watkins, Maren
  ;  APPLICANT:  Jones, Robert M.
  ;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
  ;  FILE REFERENCE: 2314-224-II
  ;  CURRENT APPLICATION NUMBER: US/10/092,367
  ;  CURRENT FILING DATE:  2002-03-07
  ;  PRIOR APPLICATION NUMBER: US 60/273,639
  ;  PRIOR FILING DATE: 2001-03-07
  ;  NUMBER OF SEQ ID NOS: 196
  ;  SOFTWARE: PatentIn version 3.0
  ; SEQ ID NO 135
  ;   LENGTH: 17
  ;   TYPE: PRT
  ;   ORGANISM: Conus betulinus
  ;   FEATURE:
  ;   NAME/KEY: PEPTIDE
  ;   LOCATION: (1)..(17)
  ;   OTHER INFORMATION: Xaa is Glu or gamma-carboxy-Glu
  US-10-092-367-135

    Query Match             75.8%;  Score 50;  DB 12;  Length 17;
    Best Local Similarity   88.2%;  Pred. No. 0.62;
    Matches   15;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

  Qy          1 GGGXVRXSAXTLHXITX 17
                || |||||||||||||:
  Db          1 GGXXVRXSAXTLHXITP 17


  RESULT 12
  US-10-092-367-61
  ; Sequence 61, Application US/10092367
  ; Publication No. US20030065138A1
  ; GENERAL INFORMATION:
  ;  APPLICANT: University of Utah Research Foundation
  ;  APPLICANT:  Cognetix, Inc.
  ;  APPLICANT:  Olivera, Baldomero M
  ;  APPLICANT:  McIntosh, J. Michael
  ;  APPLICANT:  Garrett, James E.
  ;  APPLICANT:  Walker, Craig S.
  ;  APPLICANT:  Watkins, Maren
  ;  APPLICANT:  Jones, Robert M.
  ;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
  ;  FILE REFERENCE: 2314-224-II
  ;  CURRENT APPLICATION NUMBER: US/10/092,367
  ;  CURRENT FILING DATE:  2002-03-07
  ;  PRIOR APPLICATION NUMBER: US 60/273,639
  ;  PRIOR FILING DATE: 2001-03-07
  ;  NUMBER OF SEQ ID NOS: 196
  ;  SOFTWARE: PatentIn version 3.0
  ; SEQ ID NO 61
  ;   LENGTH: 95
  ;   TYPE: PRT
  ;   ORGANISM: Conus betulinus
  US-10-092-367-61

    Query Match             75.8%;  Score 50;  DB 12;  Length 95;
    Best Local Similarity   56.2%;  Pred. No. 4.6;
    Matches    9;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

  Qy          2 GGXVRXSAXTLHXITX 17
                | :||:||:|||::|:
  Db         80 GEEVRESAETLHELTP 95


  RESULT 13
  US-10-092-367-2
  ; Sequence 2, Application US/10092367
  ; Publication No. US20030065138A1
  ; GENERAL INFORMATION:
  ;  APPLICANT: University of Utah Research Foundation
  ;  APPLICANT:  Cognetix, Inc.
  ;  APPLICANT:  Olivera, Baldomero M
  ;  APPLICANT:  McIntosh, J. Michael
  ;  APPLICANT:  Garrett, James E.
  ;  APPLICANT:  Walker, Craig S.
  ;  APPLICANT:  Watkins, Maren
  ;  APPLICANT:  Jones, Robert M.
  ;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
  ;  FILE REFERENCE: 2314-224-II
  ;  CURRENT APPLICATION NUMBER: US/10/092,367
```

; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
; OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-2

Query Match 72.7%; Score 48; DB 12; Length 17;
Best Local Similarity 88.2%; Pred. No. 1.3;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITX 17
|| |||||||||||:||
Db 1 GGXXVRXSAXTLHXLTX 17

RESULT 14
US-10-092-367-62
; Sequence 62, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Xaa at residues 3, 4, 7, 10 and 14 is Glu or gamma-carboxy-Glu; X
; OTHER INFORMATION: aa at residue 17 is Pro or hydroxy-Pro
US-10-092-367-62

Query Match 72.7%; Score 48; DB 12; Length 17;
Best Local Similarity 88.2%; Pred. No. 1.3;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITX 17
|| |||||||||||:||
Db 1 GGXXVRXSAXTLHXLTX 17

RESULT 15
US-10-092-367-134
; Sequence 134, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 134
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Conus betulinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Xaa is Glu or gamma-carboxy-Glu
US-10-092-367-134

Query Match 72.7%; Score 48; DB 12; Length 17;
Best Local Similarity 82.4%; Pred. No. 1.3;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXITX 17
|| |||||||||||:|:
Db 1 GGXXVRXSAXTLHXLTP 17

Search completed: June 2, 2004, 18:15:57
Job time : 15.3643 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:10:29 ; Search time 4.6124 Seconds
(without alignments)
354.534 Million cell updates/sec

Title: US-10-092-367-6
Perfect score: 66
Sequence: 1 GGGXVRXSAXTLHXITX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	72.7	1417	2	H83132	probable sensor/re
2	46	69.7	343	2	G91161	RNA 3'-terminal ph
3	46	69.7	403	2	T40473	hypothetical prote
4	45	68.2	146	2	G65137	hypothetical 15.4
5	44	66.7	209	2	S13179	transforming prote
6	44	66.7	217	2	H70631	hypothetical prote
7	44	66.7	792	2	A84308	chloride channel [
8	43	65.2	186	1	TVDORS	transforming prote
9	43	65.2	189	1	TVDORA	transforming prote
10	43	65.2	189	2	S33796	ras protein homolo
11	43	65.2	191	2	JC6328	Ras2 protein - sli
12	43	65.2	191	2	S58220	transforming prote
13	43	65.2	192	2	S55022	transforming prote
14	43	65.2	192	2	S32042	GTP-binding protei
15	43	65.2	193	2	S38362	Ppras2 protein - s
16	43	65.2	195	1	TVFFR	transforming prote
17	43	65.2	203	1	TVHUC2	GTP-binding protei
18	43	65.2	203	2	A36365	transforming prote
19	43	65.2	206	2	C36365	transforming prote
20	43	65.2	215	2	JN0562	hypothetical 24K p
21	43	65.2	217	1	TVWYRS	transforming prote
22	43	65.2	218	1	TVHURR	transforming prote
23	43	65.2	231	2	T32953	hypothetical prote
24	43	65.2	309	1	TVBYR1	GTP-binding protei
25	43	65.2	345	2	C90416	hypothetical prote
26	43	65.2	814	2	T30950	hypothetical prote
27	43	65.2	2712	2	T30949	hypothetical prote
28	42	63.6	118	2	AH3454	hypothetical prote
29	42	63.6	273	2	T35153	hypothetical prote
30	42	63.6	544	2	E75569	probable aminotran
31	42	63.6	650	2	T23175	hypothetical prote
32	42	63.6	654	2	A57785	finger protein ZNF
33	41	62.1	183	1	TVHUR2	transforming prote
34	41	62.1	183	2	S03180	transforming prote
35	41	62.1	268	2	A86988	conserved hypothet
36	41	62.1	269	2	S72590	hypothetical prote
37	41	62.1	390	2	F96723	hypothetical prote
38	41	62.1	673	2	B86437	F28K20.8 protein -
39	41	62.1	1508	2	T27828	hypothetical prote
40	41	62.1	1519	2	T27829	hypothetical prote
41	40.5	61.4	805	2	A56199	transcription fact
42	40	60.6	159	2	T46857	molybdenum cofacto
43	40	60.6	292	2	G72530	probable nicotine
44	40	60.6	321	2	S04253	quinate 5-dehydrog
45	40	60.6	328	2	B69429	conserved hypothet

ALIGNMENTS

RESULT 1
H83132
probable sensor/response regulator hybrid PA4112 [imported] - Pseudomonas aeruginosa (st
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: H83132
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83132
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1417 <STO>
A;Cross-references: GB:AE004827; GB:AE004091; NID:g9950306; PIDN:AAG07499.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4112

Query Match 72.7%; Score 48; DB 2; Length 1417;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              | |:|::||:|||:|
Db       1336 GEGDVQGSAATLHTI 1350
```

RESULT 2
G91161
RNA 3'-terminal phosphate cyclase [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: G91161
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genon
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G91161
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-343 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37686.1; PID:g13363737; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4263
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0025

Query Match 69.7%; Score 46; DB 2; Length 343;

Best Local Similarity 47.1%; Pred. No. 11;
Matches 8; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              ||| :: :|| :: | :|| :
Db         15 GGGQIMRSALSLSMITG 31
```

RESULT 3
T40473
hypothetical protein SPBC4B4.01c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C;Accession: T40473
R;Beck, A.; Reinhardt, R.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z21932
A;Accession: T40473
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-403 <BEC>
A;Cross-references: EMBL:AL023706; PIDN:CAA19281.1; GSPDB:GN00067; SPDB:SPBC4B4.01c
A;Experimental source: strain 972h-; cosmid c4B4
C;Genetics:
A;Gene: SPDB:SPBC4B4.01c
A;Map position: 2
A;Introns: 36/1; 65/3
C;Superfamily: Saccharomyces hypothetical protein YDR531w

Query Match 69.7%; Score 46; DB 2; Length 403;
Best Local Similarity 35.3%; Pred. No. 14;
Matches 6; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              || ::|:  :|:|::|:
Db        343 GGSFIRNHVQTMHTLTY 359
```

RESULT 4
G65137
hypothetical 15.4 kD protein in malT-glpR intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: G65137
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65137
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-146 <BLAT>
A;Cross-references: GB:AE000418; GB:U00096; NID:g2367222; PIDN:AAC76445.1; PID:g1789826;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yhgK

Query Match 68.2%; Score 45; DB 2; Length 146;
Best Local Similarity 47.1%; Pred. No. 6.6;
Matches 8; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              ||| :: :|| :: | :|| :
Db         15 GGGQILRSALSLSMITG 31
```

RESULT 5
S13179
transforming protein (ras) - Geodia cydonium
C;Species: Geodia cydonium
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S13179
R;Robitzki, A.; Schroeder, H.C.; Ugarkovic, D.; Kuchino, Y.; Kurelec, B.; Gamulin, V.; Mu
Eur. J. Biochem. 192, 499-506, 1990
A;Title: Regulated expression and phosphorylation of the 23-26-kDa ras protein in the spo
A;Reference number: S13179; MUID:91006138; PMID:2209606
A;Accession: S13179
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-209 <ROB>
A;Note: based on the evidence for Gln-tRNA, the authors translated the codon TAG as Gln;
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;140-143/Region: GTP-binding NKXD motif
F;168-170/Region: GTP-binding SAK/L motif
F;16,17,58,140,141,143,168/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

Query Match 66.7%; Score 44; DB 2; Length 209;
Best Local Similarity 53.3%; Pred. No. 14;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              ||| :| :|| :|| ::
Db         10 GGGLVGKSALTLQLV 24
```

RESULT 6
H70631
hypothetical protein Rv0434 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Jun-2003
C;Accession: H70631
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome s
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70631
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-217 <COL>
A;Cross-references: GB:Z84724; GB:AL123456; NID:g3261708; PIDN:CAB06574.1; PID:g1817700
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0434
C;Superfamily: uncharacterized protein, N-terminal domain of Lon protease homolog

Query Match 66.7%; Score 44; DB 2; Length 217;
Best Local Similarity 47.1%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXITX 17
              ||| : |:  :|| :||:
Db         60 GGGDTRCDVGTLARITE 76
```

RESULT 7
A84308
chloride channel [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: A84308
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lia
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84308
A;Status: preliminary
A;Molecule type: DNA

```
A;Residues: 1-792 <STO>
A;Cross-references: GB:AE004437; NID:g10581031; PIDN:AAG19829.1; GSPDB:GN00138
C;Genetics:
A;Gene: clc

  Query Match             66.7%;  Score 44;  DB 2;  Length 792;
  Best Local Similarity   47.1%;  Pred. No. 61;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:  :||: ||:| :
Db         57 GGGLAVVSAYNLHRIAH 73


RESULT 8
FVDORS
transforming protein ras - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 19-Jan-2001
C;Accession: A01371
R;Reymond, C.D.; Gomer, R.H.; Mehdy, M.C.; Firtel, R.A.
Cell 39, 141-148, 1984
A;Title: Developmental regulation of a Dictyostelium gene encoding a protein homologous
A;Reference number: A01371; MUID:85024887; PMID:6091907
A;Accession: A01371
A;Molecule type: DNA
A;Residues: 1-186 <REY>
A;Cross-references: GB:K02114; NID:g167864; PIDN:AAA33243.1; PID:g167865
C;Genetics:
A;Gene: ras
A;Introns: 25/3; 30/1; 47/1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane protein; methylated carboxyl end; nucleot
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;145-147/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,145/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F;183/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;183/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

  Query Match             65.2%;  Score 43;  DB 1;  Length 186;
  Best Local Similarity   53.3%;  Pred. No. 19;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 9
FVDORA
transforming protein rasG - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 19-Jan-2001
C;Accession: A31456; S21090; S22129
R;Robbins, S.M.; Williams, J.G.; Jermyn, K.A.; Spiegelman, G.B.; Weeks, G.
Proc. Natl. Acad. Sci. U.S.A. 86, 938-942, 1989
A;Title: Growing and developing Dictyostelium cells express different ras genes.
A;Reference number: A31456; MUID:89128893; PMID:2644652
A;Accession: A31456
A;Molecule type: mRNA
A;Residues: 1-189 <ROB1>
A;Cross-references: GB:J04160; NID:g167866; PIDN:AAA33244.1; PID:g167867
R;Robbins, S.M.; Williams, J.G.; Spiegelman, G.B.; Weeks, G.
Biochim. Biophys. Acta 1130, 85-89, 1992
A;Title: Cloning and characterization of the Dictyostelium discoideum rasG genomic seque
A;Reference number: S21090; MUID:92182019; PMID:1339294
A;Accession: S21090
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-189 <ROB2>
A;Cross-references: EMBL:Z11533; NID:g7342; PIDN:CAA77632.1; PID:g7343
C;Genetics:
A;Gene: rasG
A;Introns: 25/3
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane protein; methylated carboxyl end; nucleoti
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat
F;186/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;186/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

  Query Match             65.2%;  Score 43;  DB 1;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 19;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 10
S33796
ras protein homolog - slime mold (Physarum polycephalum)
C;Species: Physarum polycephalum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S33796
R;Kozlowski, P.; Fronk, J.; Toczko, K.
Biochim. Biophys. Acta 1173, 357-359, 1993
A;Title: Identification of a ras gene in the slime mold Physarum polycephalum.
A;Reference number: S33796; MUID:93305735; PMID:8318547
A;Accession: S33796
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-189 <KOZ>
A;Cross-references: GB:L10344; GB:Z21495; NID:g1478117; PIDN:AAB05646.1; PID:g310554
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

  Query Match             65.2%;  Score 43;  DB 2;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 19;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 11
JC6328
Ras2 protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 21-May-1998 #sequence_revision 29-May-1998 #text_change 19-Jan-2001
C;Accession: JC6328
R;van Es, S.; Kooistra, R.A.; Schaap, P.
Gene 187, 93-97, 1997
A;Title: Two ras genes in Dictyostelium minutum show high sequence homology, but differer
A;Reference number: JC6304; MUID:97225801; PMID:9073071
A;Accession: JC6328
A;Molecule type: DNA
A;Residues: 1-191 <VAN>
C;Comment: This protein is expressed during the entire course of development and is not r
C;Genetics:
A;Gene: ras2
A;Introns: 25/2; 30/1; 65/2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
```

F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17,57-62,115-118,144-148/Domain: GTP-binding #status predicted <GTB>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 65.2%; Score 43; DB 2; Length 191;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24
```

RESULT 12
S58220
transforming protein ras-2 - Dictyostelium minutum
C;Species: Dictyostelium minutum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S58220
R;van Es, S.; Kooistra, R.A.; Schaap, P.
submitted to the EMBL Data Library, July 1995
A;Description: Two ras genes in Dictyostelium minutum show high sequence homology, but d
A;Reference number: S58220
A;Accession: S58220
A;Molecule type: DNA
A;Residues: 1-191 <VAN>
A;Cross-references: EMBL:X89037; NID:g929568; PIDN:CAA61434.1; PID:g929569
A;Experimental source: strain 71-2
C;Genetics:
A;Gene: ras2
A;Introns: 25/2; 30/1; 55/2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane protein; methylated carboxyl end; nucleot
F;4-119/Domain: translation elongation factor Tu homology <ETU>
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;116-119/Region: GTP-binding NKXD motif
F;146-148/Region: GTP-binding SAK/L motif
F;16,17,35,116,117,119,146/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F;188/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;188/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 65.2%; Score 43; DB 2; Length 191;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24
```

RESULT 13
S55022
transforming protein ras2 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 19-Jan-2001
C;Accession: S55022; S12083
R;Harrison, S.D.; Solomon, N.; Rubin, G.M.
Genetics 139, 1701-1709, 1995
A;Title: A genetic analysis of the 63E-64A genomic region of Drosophila melanogaster: id
A;Reference number: S55020; MUID:95309683; PMID:7789770
A;Accession: S55022
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-192 <HAR>
A;Cross-references: EMBL:U15967; NID:g639707; PIDN:AAB60243.1; PID:g639710
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R;Cohen, N.; Salzberg, A.; Lev, Z.
Oncogene 3, 137-142, 1988
A;Title: A bidirectional promoter is regulating the Drosophila ras2 gene.
A;Reference number: S12083; MUID:88319648; PMID:3412773
A;Accession: S12083
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-27,'VS' <COH>
A;Cross-references: EMBL:X07255; NID:g8402; PIDN:CAA30242.1; PID:g8403
C;Genetics:
A;Gene: ras2
A;Cross-references: FlyBase:FBgn0003206
A;Introns: 27/3; 57/1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop; transforming protein
F;6-121/Domain: translation elongation factor Tu homology <ETU>
F;12-19/Region: nucleotide-binding motif A (P-loop)
F;118-121/Region: GTP-binding NKXD motif
F;148-150/Region: GTP-binding SAK/L motif
F;18,19,37,118,119,121,148/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

Query Match 65.2%; Score 43; DB 2; Length 192;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSAITIQFI 26
```

RESULT 14
S32042
GTP-binding protein ras2 - Hydra magnipapillata
C;Species: Hydra magnipapillata
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 19-Jan-2001
C;Accession: JC4573; S32042
R;Bosch, T.C.G.; Benitez, E.; Gellner, K.; Praetzel, G.; Salgado, L.M.
Gene 167, 191-195, 1995
A;Title: Cloning of a ras-related gene from Hydra which responds to head-specific signals
A;Reference number: JC4573; MUID:96144273; PMID:8566776
A;Accession: JC4573
A;Molecule type: mRNA
A;Residues: 1-192 <BOS>
A;Cross-references: EMBL:X70839; NID:g11139; PIDN:CAA50187.1; PID:g11140
A;Experimental source: epithelial cell
C;Comment: This protein is a member of ras protein family, and a key component in recepto
. This protein is highly sensitive to head-specific signals and plays a critical role in
C;Genetics:
A;Gene: ras2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; methylated carboxyl end; nucleotide binding; P-loop
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;37-45/Region: effector
F;58-63/Region: nucleotide-binding motif B
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat
F;189/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted
F;189/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 65.2%; Score 43; DB 2; Length 192;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         15 GGGGVGKSALTIQFI 29
```

RESULT 15
S38362
Ppras2 protein - slime mold (Physarum polycephalum)
C;Species: Physarum polycephalum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-Jan-2001
C;Accession: S38362
R;Kozlowski, P.; Tymowska, Z.; Toczko, K.

Biochim. Biophys. Acta 1174, 299-302, 1993
A;Title: Nucleotide and predicted amino acid sequence of a new member of the ras gene fa
A;Reference number: S38362; MUID:93385161; PMID:8373809
A;Accession: S38362
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-193 <KOZ>
A;Cross-references: GB:L14275; NID:g404808; PIDN:AAC37179.1; PID:g404809
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;6-121/Domain: translation elongation factor Tu homology <ETU>
F;12-19/Region: nucleotide-binding motif A (P-loop)
F;118-121/Region: GTP-binding NKXD motif
F;148-150/Region: GTP-binding SAK/L motif
F;18,19,37,118,119,121,148/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 65.2%; Score 43; DB 2; Length 193;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGXVRXSAXTLHXI 15
||||:| :||:|: :|
Db 12 GGGGVGKSALTIQLI 26

Search completed: June 2, 2004, 18:13:07
Job time : 12.6124 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:06:18 ; Search time 3.16279 Seconds
(without alignments)
279.877 Million cell updates/sec

Title: US-10-092-367-6
Perfect score: 66
Sequence: 1 GGGXVRXSAXTLHXITX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	69.7	342	1	RTCA_ECO57	P58127 escherichia
2	45	68.2	338	1	RTCA_ECOLI	P46849 escherichia
3	45	68.2	339	1	RTCA_SALTY	Q8zli0 salmonella
4	44	66.7	187	1	DEF_CHLTE	Q8kcg7 chlorobium
5	44	66.7	209	1	RAS_GEOCY	P24498 geodia cydo
6	44	66.7	338	1	RTCA_ECOL6	Q8fcs8 escherichia
7	44	66.7	347	1	RTCA_RALSO	Q8y2v6 ralstonia s
8	43	65.2	187	1	RASD_DICDI	P03967 dictyosteli
9	43	65.2	189	1	RAS1_PHYPO	P34729 physarum po
10	43	65.2	189	1	RASG_DICDI	P15064 dictyosteli
11	43	65.2	192	1	RAS2_DROME	P04388 drosophila
12	43	65.2	192	1	RAS2_HYDMA	P38976 hydra magni
13	43	65.2	193	1	RAS2_PHYPO	P34726 physarum po
14	43	65.2	197	1	RASB_DICDI	P32252 dictyosteli
15	43	65.2	203	1	RAS1_RHIRA	P22278 rhizomucor
16	43	65.2	204	1	RRA2_HUMAN	P17082 homo sapien
17	43	65.2	205	1	RAS3_RHIRA	P22280 rhizomucor
18	43	65.2	215	1	RASL_COPCI	Q05058 coprinus ci
19	43	65.2	216	1	RAS_CRYNE	O74650 cryptococcu
20	43	65.2	217	1	RAS_LENED	P28775 lentinula e
21	43	65.2	218	1	RRAS_HUMAN	P10301 homo sapien
22	43	65.2	218	1	RRAS_MOUSE	P10833 mus musculu
23	43	65.2	290	1	RAS1_CANAL	Q9uqx7 candida alb
24	43	65.2	309	1	RAS1_YEAST	P01119 saccharomyc
25	43	65.2	337	1	RTCA_SULSO	Q97w04 sulfolobus
26	42	63.6	342	1	RTCA_PYRFU	Q8u0n7 pyrococcus
27	42	63.6	460	1	NIFN_RHILO	Q98ap3 rhizobium l
28	42	63.6	654	1	Z133_HUMAN	P52736 homo sapien
29	42	63.6	715	1	BBS2_BRARE	Q98sp7 brachydanio
30	42	63.6	751	1	Z337_HUMAN	Q9y3m9 homo sapien
31	41	62.1	183	1	RAP2_HUMAN	P10114 homo sapien
32	41	62.1	183	1	RAP3_HUMAN	P17964 homo sapien
33	41	62.1	268	1	RECO_MYCLE	Q9ccn0 mycobacteri
34	41	62.1	339	1	RTCA_SULTO	Q974u1 sulfolobus
35	41	62.1	512	1	DNLI_STRCO	Q9fcb1 streptomyce
36	41	62.1	779	1	EFG2_HUMAN	Q969s9 homo sapien
37	40.5	61.4	805	1	E2F_DROME	Q27368 drosophila
38	40	60.6	159	1	MOAC_RHOSH	Q9zfa6 rhodobacter
39	40	60.6	210	1	C1IB_HUMAN	Q96iu4 homo sapien
40	40	60.6	321	1	DHQA_NEUCR	P11635 neurospora
41	40	60.6	328	1	RTCA_ARCFU	O28837 archaeoglob
42	40	60.6	355	1	RTCA_METAC	Q8th85 methanosarc
43	40	60.6	473	1	PCC6_MYCTU	Q10506 mycobacteri
44	40	60.6	1586	1	ARO1_EMENI	P07547 e pentafunc
45	40	60.6	1723	1	PM20_CHLPN	Q9z812 chlamydia p

ALIGNMENTS

RESULT 1
RTCA_ECO57
ID RTCA_ECO57 STANDARD; PRT; 342 AA.
AC P58127;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-phosphate
DE cyclase) (RNA cyclase).
GN RTCA OR Z4778 OR ECS4263.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC cyclic phosphodiester at the end of RNA. The mechanism of action
CC of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC the phosphorus in the diester linkage to produce the cyclic end
CC product. The biological role of this enzyme is unknown but it is
CC likely to function in some aspects of cellular RNA processing (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC Subfamily 1.
CC ----------

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE005564; AAG58524.1; -.
DR   EMBL; AP002565; BAB37686.1; ALT_INIT.
DR   HSSP; P46849; 1QMH.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Ligase; Complete proteome.
FT   ACT_SITE    308    308       BY SIMILARITY.
FT   DISULFID    307    307       INTERCHAIN (BY SIMILARITY).
SQ   SEQUENCE   342 AA;  36332 MW;  783FE7FAD7160846 CRC64;

  Query Match             69.7%;  Score 46;  DB 1;  Length 342;
  Best Local Similarity   47.1%;  Pred. No. 3.1;
  Matches    8;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: :||::| :||:
Db         14 GGGQIMRSALSLSMITG 30


RESULT 2
RTCA_ECOLI
ID   RTCA_ECOLI     STANDARD;      PRT;   338 AA.
AC   P46849; P46848; Q47349;
DT   01-NOV-1995 (Rel. 32, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-phosphate
DE   cyclase) (RNA cyclase).
GN   RTCA OR B3419/B3420.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE OF 149-339 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=86275993; PubMed=3015733;
RA   Cole S.T., Raibaud O.;
RT   "The nucleotide sequence of the malT gene encoding the positive
RT   regulator of the Escherichia coli maltose regulon.";
RL   Gene 42:201-208(1986).
RN   [3]
RP   REVISION, AND CHARACTERIZATION.
RX   MEDLINE=97327572; PubMed=9184239;
RA   Genschik P., Billy E., Swianiewicz M., Filipowicz W.;
RT   "The human RNA 3'-terminal phosphate cyclase is a member of a new
RT   family of proteins conserved in Eucarya, Bacteria and Archaea.";
RL   EMBO J. 16:2955-2967(1997).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=98411361; PubMed=9738023;
RA   Genschik P., Drabikowski K., Filipowicz W.;
RT   "Characterization of the Escherichia coli RNA 3'-terminal phosphate
RT   cyclase and its sigma54-regulated operon.";
RL   J. Biol. Chem. 273:25516-25526(1998).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RC   STRAIN=K12;
RX   MEDLINE=20139688; PubMed=10673421;
RA   Palm G.J., Billy E., Filipowicz W., Wlodawer A.;
RT   "Crystal structure of RNA 3'-terminal phosphate cyclase, a ubiquitous
RT   enzyme with unusual topology.";
RL   Structure 8:13-23(2000).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing.
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 122 that produces two separate ORFs.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18997; AAA58218.1; ALT_FRAME.
DR   EMBL; U18997; AAA58217.1; ALT_FRAME.
DR   EMBL; AE000418; AAC76445.1; ALT_FRAME.
DR   EMBL; AE000418; AAC76444.1; ALT_FRAME.
DR   EMBL; M13585; AAA83889.1; -.
DR   PDB; 1QMH; 11-JAN-00.
DR   PDB; 1QMI; 11-JAN-00.
DR   EcoGene; EG12938; rtcA.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Ligase; 3D-structure; Complete proteome.
FT   ACT_SITE    308    308       PROBABLE.
FT   DISULFID    307    307       INTERCHAIN.
FT   STRAND        5      8
FT   TURN          9     10
FT   TURN         12     13
FT   HELIX        16     29
FT   STRAND       33     36
FT   TURN         38     41
FT   STRAND       42     42
FT   HELIX        49     62
FT   TURN         63     63
FT   STRAND       65     67
FT   TURN         71     72
FT   STRAND       76     79
FT   STRAND       86     97
FT   HELIX        98    109
FT   TURN        110    111
FT   STRAND      116    123
FT   STRAND      125    126
FT   TURN        127    128
FT   STRAND      129    129
FT   TURN        132    132
FT   HELIX       133    137
FT   TURN        138    138
FT   HELIX       139    145
FT   TURN        146    147
```

```
FT   STRAND      149    156
FT   STRAND      159    159
FT   TURN        160    161
FT   STRAND      165    172
FT   STRAND      181    181
FT   STRAND      184    184
FT   STRAND      188    198
FT   HELIX       202    215
FT   STRAND      220    226
FT   HELIX       228    230
FT   STRAND      233    242
FT   STRAND      246    252
FT   TURN        255    256
FT   HELIX       259    275
FT   STRAND      278    278
FT   HELIX       282    295
FT   TURN        296    296
FT   STRAND      299    302
FT   HELIX       307    319
FT   STRAND      325    328
FT   STRAND      333    336
SQ   SEQUENCE   338 AA;  35903 MW;  3450201CB8E40CE7 CRC64;

  Query Match             68.2%;  Score 45;  DB 1;  Length 338;
  Best Local Similarity   47.1%;  Pred. No. 4.6;
  Matches    8;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: :||::| :||:
Db         14 GGGQILRSALSLSMITG 30


RESULT 3
RTCA_SALTY
ID   RTCA_SALTY     STANDARD;      PRT;   339 AA.
AC   Q8ZLI0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-phosphate
DE   cyclase) (RNA cyclase).
GN   RTCA OR STM3518.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008862; AAL22380.1; -.
DR   StyGene; SG?????; rtcA.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; FALSE_NEG.
KW   Ligase; Complete proteome.
FT   ACT_SITE    308    308       BY SIMILARITY.
SQ   SEQUENCE   339 AA;  35457 MW;  182667CD81E31125 CRC64;

  Query Match             68.2%;  Score 45;  DB 1;  Length 339;
  Best Local Similarity   47.1%;  Pred. No. 4.6;
  Matches    8;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: :||::| :||:
Db         14 GGGQILRSALSLSMITG 30


RESULT 4
DEF_CHLTE
ID   DEF_CHLTE      STANDARD;      PRT;   187 AA.
AC   Q8KCG7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN   DEF OR CT1454.
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC   -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC       newly synthesized proteins. Requires at least a dipeptide for an
CC       efficient rate of reaction. N-terminal L-methionine is a
CC       prerequisite for activity but the enzyme has broad specificity at
CC       other positions (By similarity).
CC   -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC       methionyl peptide.
CC   -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC   -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; AE012902; AAM72682.1; -.
DR   TIGR; CT1454; -.
DR   HAMAP; MF_00163; -; 1.
DR   InterPro; IPR000181; Pep_deformylase.
DR   Pfam; PF01327; Pep_deformylase; 1.
DR   PRINTS; PR01576; PDEFORMYLASE.
DR   ProDom; PD003844; Pep_deformylase; 1.
DR   TIGRFAMs; TIGR00079; pept_deformyl; 1.
KW   Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT   ACT_SITE    137    137       BY SIMILARITY.
FT   METAL        94     94       IRON (BY SIMILARITY).
FT   METAL       136    136       IRON (BY SIMILARITY).
FT   METAL       140    140       IRON (BY SIMILARITY).
SQ   SEQUENCE   187 AA;  20909 MW;  1E16EA5077AFC296 CRC64;

  Query Match             66.7%;  Score 44;  DB 1;  Length 187;
  Best Local Similarity   61.5%;  Pred. No. 3.5;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGXVRXSAXTLHX 14
              | :||:||:|||:
Db        102 GNVVRPSAITLHY 114


RESULT 5
RAS_GEOCY
ID   RAS_GEOCY      STANDARD;      PRT;   209 AA.
AC   P24498;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein.
OS   Geodia cydonium (Sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Astrophorida; Geodiidae; Geodia.
OX   NCBI_TaxID=6047;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91006138; PubMed=2209606;
RA   Robitzki A., Schroeder H.C., Ugarkovic D., Kuchino Y., Kurelec B.,
RA   Gamulin V., Mueller W.E.G.;
RT   "Regulated expression and phosphorylation of the 23-26-kDa ras
RT   protein in the sponge Geodia cydonium.";
RL   Eur. J. Biochem. 192:499-506(1990).
CC   -!- FUNCTION: This protein is activated by the insulin/insulin
CC       (insulin-like)-receptor system. This transition enables the ras
CC       protein to interact with the lectin-receptor/lectin complex, a
CC       process which ultimately lead to an initiation of an intra-
CC       cellular signal-transduction chain.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- PTM: PHOSPHORYLATED IN THE PRESENCE OF INSULIN.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30929; -; NOT_ANNOTATED_CDS.
DR   PIR; S13179; S13179.
DR   HSSP; P01112; 1PLJ.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
KW   GTP-binding; Prenylation; Lipoprotein; Phosphorylation.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      79     83       GTP (BY SIMILARITY).
FT   NP_BIND     140    143       GTP (BY SIMILARITY).
FT   DOMAIN       55     63       EFFECTOR REGION (BY SIMILARITY).
FT   MOD_RES      58     58       PHOSPHORYLATION (POTENTIAL).
FT   LIPID       206    206       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   209 AA;  23854 MW;  C544C43102C8323D CRC64;

  Query Match             66.7%;  Score 44;  DB 1;  Length 209;
  Best Local Similarity   53.3%;  Pred. No. 4;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|| ::
Db         10 GGGLVGKSALTLQLV 24


RESULT 6
RTCA_ECOL6
ID   RTCA_ECOL6      STANDARD;      PRT;   338 AA.
AC   Q8FCS8;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-
DE   phosphate cyclase) (RNA cyclase).
GN   RTCA OR C4197.
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE016768; AAN82635.1; ALT_INIT.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Ligase; Complete proteome.
FT   ACT_SITE    308    308       BY SIMILARITY.
```

```
FT   DISULFID    307    307       INTERCHAIN (BY SIMILARITY).
SQ   SEQUENCE   338 AA;  35949 MW;  38075DF4CA2CBC33 CRC64;

  Query Match             66.7%;  Score 44;  DB 1;  Length 338;
  Best Local Similarity   47.1%;  Pred. No. 6.9;
  Matches    8;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: :||::| :||:
Db         14 GGGQILRSALSLPMITG 30


RESULT 7
RTCA_RALSO
ID   RTCA_RALSO     STANDARD;      PRT;   347 AA.
AC   Q8Y2V6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-
DE   phosphate cyclase) (RNA cyclase).
GN   RTCA OR RSC0226 OR RS00658.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
CC   -!- FUNCTION: Catalyzes the conversion of 3'-phosphate to a 2',3'-
CC       cyclic phosphodiester at the end of RNA. The mechanism of action
CC       of the enzyme occurs in 3 steps: (A) adenylation of the enzyme by
CC       ATP; (B) the enzyme acts on RNA-N3'P to produce RNA-N3'PP5'A; (C)
CC       a non catalytic nucleophilic attack by the adjacent 2'hydroxyl on
CC       the phosphorus in the diester linkage to produce the cyclic end
CC       product. The biological role of this enzyme is unknown but it is
CC       likely to function in some aspects of cellular RNA processing (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +
CC       diphosphate + RNA terminal-2',3'-cyclic-phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the RNA 3'-terminal cyclase family.
CC       Subfamily 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL646058; CAD13754.1; -.
DR   HAMAP; MF_00200; -; 1.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Ligase; Complete proteome.
FT   ACT_SITE    315    315       BY SIMILARITY.
SQ   SEQUENCE   347 AA;  35970 MW;  913E69C707B70524 CRC64;

  Query Match             66.7%;  Score 44;  DB 1;  Length 347;
  Best Local Similarity   41.2%;  Pred. No. 7.1;
  Matches    7;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: ::|:|| ::|:
Db         20 GGGQILRTALTLSMLTG 36


RESULT 8
RASD_DICDI
ID   RASD_DICDI     STANDARD;      PRT;   187 AA.
AC   P03967;
DT   23-OCT-1986 (Rel. 02, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein rasD (Transforming protein P23).
GN   RASD OR RASA OR RAS.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=85024887; PubMed=6091907;
RA   Reymond C.D., Gomer R.H., Mehdy M.C., Firtel R.A.;
RT   "Developmental regulation of a Dictyostelium gene encoding a protein
RT   homologous to mammalian ras protein.";
RL   Cell 39:141-148(1984).
RN   [2]
RP   REVISIONS.
RC   STRAIN=AX3;
RX   MEDLINE=91115102; PubMed=1703508;
RA   Esch R.K., Firtel R.A.;
RT   "cAMP and cell sorting control the spatial expression of a
RT   developmentally essential cell-type-specific ras gene in
RT   Dictyostelium.";
RL   Genes Dev. 5:9-21(1991).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- DEVELOPMENTAL STAGE: Expressed at a low level in vegetative cells;
CC       not expressed between the onset of development and aggregation,
CC       and is then re-expressed in the multicellular aggregate stages.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02114; AAA33243.1; -.
DR   EMBL; Z11804; CAA77848.1; -.
DR   PIR; A01371; TVDORS.
DR   HSSP; P01112; 1PLK.
DR   DictyBase; DDB0001711; rasD.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      57     61       GTP (BY SIMILARITY).
FT   NP_BIND     116    119       GTP (BY SIMILARITY).
FT   DOMAIN       32     40       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       184    184       S-geranylgeranyl cysteine
```

```
FT                                (By similarity).
FT   CONFLICT    137    143       GFNCPFM -> DSLSFH (IN REF. 1).
SQ   SEQUENCE   187 AA;  21202 MW;  7F526253B8316678 CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 187;
  Best Local Similarity   53.3%;  Pred. No. 5.3;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 9
RAS1_PHYPO
ID   RAS1_PHYPO     STANDARD;      PRT;   189 AA.
AC   P34729;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein 1.
GN   RAS1 OR RAS-1.
OS   Physarum polycephalum (Slime mold).
OC   Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
OC   Physarum.
OX   NCBI_TaxID=5791;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LU352;
RX   MEDLINE=93305735; PubMed=8318547;
RA   Kozlowski P., Fronk J., Toczko K.;
RT   "Identification of a ras gene in the slime mold Physarum
RT   polycephalum.";
RL   Biochim. Biophys. Acta 1173:357-359(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M3CVIII;
RX   MEDLINE=96186923; PubMed=8635743;
RA   Trzcińska-Danielewicz J., Kozlowski P., Toczko K.;
RT   "Cloning and genomic sequence of the Physarum polycephalum Ppras1
RT   gene, a homologue of the ras protooncogene.";
RL   Gene 169:143-144(1996).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Inner surface of plasma membrane.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10344; AAB05646.1; -.
DR   EMBL; U10905; AAB06296.1; -.
DR   PIR; S33796; S33796.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      57     61       GTP (BY SIMILARITY).
FT   NP_BIND     116    119       GTP (BY SIMILARITY).
FT   DOMAIN       32     40       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       186    186       S-geranylgeranyl cysteine
FT                                (By similarity).
```

```
SQ   SEQUENCE   189 AA;  21202 MW;  5EEC8AD372A4CB94 CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 5.3;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 10
RASG_DICDI
ID   RASG_DICDI     STANDARD;      PRT;   189 AA.
AC   P15064;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein rasG.
GN   RASG.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89128893; PubMed=2644652;
RA   Robbins S.M., Williams J.G., Jermyn K.A., Spiegelman G.B., Weeks G.;
RT   "Growing and developing Dictyostelium cells express different ras
RT   genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:938-942(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX2;
RX   MEDLINE=92182019; PubMed=1339294;
RA   Robbins S.M., Williams J.G., Spiegelman G.B., Weeks G.;
RT   "Cloning and characterization of the Dictyostelium discoideum rasG
RT   genomic sequences.";
RL   Biochim. Biophys. Acta 1130:85-89(1992).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04160; AAA33244.1; -.
DR   EMBL; Z11533; CAA77632.1; -.
DR   PIR; A31456; TVDORA.
DR   HSSP; P01112; 1PLK.
DR   DictyBase; DDB0001821; rasG.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      10     17       GTP (BY SIMILARITY).
FT   NP_BIND      57     61       GTP (BY SIMILARITY).
FT   NP_BIND     116    119       GTP (BY SIMILARITY).
FT   DOMAIN       32     40       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       186    186       S-geranylgeranyl cysteine
FT                                (By similarity).
```

```
SQ   SEQUENCE   189 AA;  21333 MW;  AFB502319C090899 CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 189;
  Best Local Similarity   53.3%;  Pred. No. 5.3;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         10 GGGGVGKSALTIQLI 24


RESULT 11
RAS2_DROME
ID   RAS2_DROME     STANDARD;      PRT;   192 AA.
AC   P04388; Q9VZH7;
DT   20-MAR-1987 (Rel. 04, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein 2.
GN   RAS64B OR RAS2.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85187987; PubMed=3921827;
RA   Mozer B., Marlor R., Parkhusrt S., Corces V.G.;
RT   "Characterization and developmental expression of a Drosophila ras
RT   oncogene.";
RL   Mol. Cell. Biol. 5:885-889(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87248071; PubMed=3110012;
RA   Brock H.W.;
RT   "Sequence and genomic structure of ras homologues Dmras85D and
RT   Dmras64B of Drosophila melanogaster.";
RL   Gene 51:129-137(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Iso-1 / Kennison;
RX   MEDLINE=95309683; PubMed=7789770;
RA   Harrison S.D., Solomon N., Rubin G.M.;
RT   "A genetic analysis of the 63E-64A genomic region of Drosophila
RT   melanogaster: identification of mutations in a replication factor C
RT   subunit.";
RL   Genetics 139:1701-1709(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [5]
RP   REVISIONS.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN   [7]
RP   SEQUENCE OF 28-192 FROM N.A.
RX   MEDLINE=84259319; PubMed=6430564;
RA   Neuman-Silberberg F.S., Schejter E., Hoffmann F.M., Shilo B.-Z.;
RT   "The Drosophila ras oncogenes: structure and nucleotide sequence.";
RL   Cell 37:1027-1033(1984).
RN   [8]
RP   SEQUENCE OF 28-192 FROM N.A.
RC   STRAIN=A1;
RX   MEDLINE=20020328; PubMed=10552039;
RA   Gasperini R., Gibson G.;
RT   "Absence of protein polymorphism in the Ras genes of Drosophila
RT   melanogaster.";
RL   J. Mol. Evol. 49:583-590(1999).
RN   [9]
RP   SEQUENCE OF 1-18 AND 44-64 FROM N.A., SPLICE SITES, AND MUTAGENESIS.
RX   MEDLINE=88255843; PubMed=2838380;
RA   Bishop J.G. III, Corces V.G.;
RT   "Expression of an activated ras gene causes developmental
RT   abnormalities in transgenic Drosophila melanogaster.";
RL   Genes Dev. 2:567-577(1988).
RN   [10]
RP   SEQUENCE OF 1-29 FROM N.A.
RX   MEDLINE=88319648; PubMed=3412773;
RA   Cohen N., Salzberg A., Lev Z.;
RT   "A bidirectional promoter is regulating the Drosophila ras2 gene.";
RL   Oncogene 3:137-142(1988).
RN   [11]
RP   CHARACTERIZATION.
RX   MEDLINE=94008534; PubMed=8404533;
RA   Salzberg A., Cohen N., Halachmi N., Kimchie Z., Lev Z.;
RT   "The Drosophila Ras2 and Rop gene pair: a dual homology with a yeast
RT   Ras-like gene and a suppressor of its loss-of-function phenotype.";
RL   Development 117:1309-1319(1993).
```

```
CC   -!- FUNCTION: May be involved in endocytic processes and/or other
CC       transport pathways mediated by vesicle trafficking. May interact
CC       functionally with ROP protein. Ras proteins bind GDP/GTP and
CC       possess intrinsic GTPase activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- DEVELOPMENTAL STAGE: A uniform expression is seen in unfertilized
CC       eggs, embryos, larvae, pupae and adult flies. Expression during
CC       embryogenesis is restricted to the CNS and the Garland cells, a
CC       small group of nephrocytes that takes up waste materials from the
CC       hemolymph by endocytosis. In post-embryonic stages, expression is
CC       seen in the larval salivary glands and the CNS, and in the adult
CC       CNS and reproductive systems.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10804; AAA99202.1; ALT_SEQ.
DR   EMBL; M10759; AAA99202.1; JOINED.
DR   EMBL; M10803; AAA99202.1; JOINED.
DR   EMBL; M16431; AAA28849.1; -.
DR   EMBL; M16124; AAA28849.1; JOINED.
DR   EMBL; M16430; AAA28849.1; JOINED.
DR   EMBL; U15967; AAB60243.1; -.
DR   EMBL; AE003480; AAF47845.2; -.
DR   EMBL; AY119135; AAM50995.1; -.
DR   EMBL; K01962; AAA28848.1; ALT_SEQ.
DR   EMBL; K01961; AAA28848.1; JOINED.
DR   EMBL; AF186651; AAF15517.1; -.
DR   EMBL; X12559; CAA31072.1; -.
DR   EMBL; X12558; CAA31071.1; ALT_INIT.
DR   EMBL; X07255; CAA30242.1; -.
DR   PIR; S55022; S55022.
DR   HSSP; P01112; 1PLK.
DR   FlyBase; FBgn0003206; Ras64B.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      12     19       GTP (BY SIMILARITY).
FT   NP_BIND      59     63       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       34     42       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       189    189       S-farnesyl cysteine (By similarity).
FT   MUTAGEN      14     14       G->V: CAUSE DEVELOPMENTAL ABNORMALITIES.
FT   CONFLICT     28     29       SY -> VS (IN REF. 10).
SQ   SEQUENCE   192 AA;  22235 MW;  3F58A3A33E8FDEBC CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 192;
  Best Local Similarity   53.3%;  Pred. No. 5.4;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSAITIQFI 26


RESULT 12
RAS2_HYDMA
ID   RAS2_HYDMA     STANDARD;      PRT;   192 AA.
AC   P38976;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein RAS2.
GN   RAS2.
OS   Hydra magnipapillata (Hydra).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6085;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=105;
RX   MEDLINE=96144273; PubMed=8566776;
RA   Bosch T.C.G., Benitez E., Gellner K., Praetzel G., Salgado L.M.;
RT   "Cloning of a ras-related gene from Hydra which responds to head-
RT   specific signals.";
RL   Gene 167:191-195(1995).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- DEVELOPMENTAL STAGE: Ras2 level drops significantly just after the
CC       head is cut. The expression goes up again after 4 to 8 hours.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70839; CAA50187.1; -.
DR   PIR; JC4573; S32042.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      62     66       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       189    189       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   192 AA;  21787 MW;  2DC2ECC18F10C709 CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 192;
  Best Local Similarity   53.3%;  Pred. No. 5.4;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         15 GGGGVGKSALTIQFI 29


RESULT 13
RAS2_PHYPO
ID   RAS2_PHYPO     STANDARD;      PRT;   193 AA.
AC   P34726;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein 2.
GN   RAS-2.
OS   Physarum polycephalum (Slime mold).
```

```
OC   Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
OC   Physarum.
OX   NCBI_TaxID=5791;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LU352;
RX   MEDLINE=93385161; PubMed=8373809;
RA   Kozlowski P., Tymowska Z., Toczko K.;
RT   "Nucleotide and predicted amino acid sequence of a new member of the
RT   ras gene family from the slime mold Physarum polycephalum.";
RL   Biochim. Biophys. Acta 1174:299-302(1993).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Inner surface of plasma membrane.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14275; AAC37179.1; -.
DR   PIR; S38362; S38362.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      12     19       GTP (BY SIMILARITY).
FT   NP_BIND      59     63       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       34     42       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       190    190       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   193 AA;  21634 MW;  4B0B33CD890EE6CD CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 193;
  Best Local Similarity   53.3%;  Pred. No. 5.5;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         12 GGGGVGKSALTIQLI 26


RESULT 14
RASB_DICDI
ID   RASB_DICDI     STANDARD;      PRT;   197 AA.
AC   P32252;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ras-like protein rasB.
GN   RASB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93205383; PubMed=8455930;
RA   Daniel J.M., Spiegelman G.B., Weeks G.;
RT   "Characterization of a third ras gene, rasB, that is expressed
RT   throughout the growth and development of Dictyostelium discoideum.";
RL   Oncogene 8:1041-1047(1993).
CC   -!- FUNCTION: Ras proteins bind GDP/GTP and possess intrinsic GTPase
CC       activity.
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96622; AAA33246.1; -.
DR   HSSP; P01112; 1PLL.
DR   DictyBase; DDB0001989; rasB.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      13     20       GTP (BY SIMILARITY).
FT   NP_BIND      60     64       GTP (BY SIMILARITY).
FT   NP_BIND     119    122       GTP (BY SIMILARITY).
FT   DOMAIN       35     43       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       194    194       S-geranylgeranyl cysteine
FT                                (By similarity).
SQ   SEQUENCE   197 AA;  22268 MW;  A3D8D3C6846BD9F4 CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 197;
  Best Local Similarity   53.3%;  Pred. No. 5.6;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         13 GGGGVGKSALTIQFI 27


RESULT 15
RAS1_RHIRA
ID   RAS1_RHIRA     STANDARD;      PRT;   203 AA.
AC   P22278;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ras-like protein 1.
GN   RAS1.
OS   Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Mucor.
OX   NCBI_TaxID=4841;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 1216B;
RX   MEDLINE=91061774; PubMed=1701021;
RA   Casale W.L., McConnell D.G., Wang S.-Y., Lee Y.-J., Linz J.E.;
RT   "Expression of a gene family in the dimorphic fungus Mucor racemosus
RT   which exhibits striking similarity to human ras genes.";
RL   Mol. Cell. Biol. 10:6654-6663(1990).
CC   -!- ENZYME REGULATION: Alternate between an inactive form bound to GDP
CC       and an active form bound to GTP. Activated by a guanine
CC       nucleotide-exchange factor (GEF) and inactivated by a GTPase-
CC       activating protein (GAP).
CC   -!- SUBCELLULAR LOCATION: Plasma membrane.
CC   -!- DEVELOPMENTAL STAGE: In all developmental stages analyzed. Its
CC       signal was more intense in sporulating mycelium.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Ras family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M55175; AAA83378.1; -.
DR   PIR; A36365; A36365.
DR   HSSP; P01112; 1PLL.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      17     24       GTP (BY SIMILARITY).
FT   NP_BIND      64     68       GTP (BY SIMILARITY).
FT   NP_BIND     123    126       GTP (BY SIMILARITY).
FT   DOMAIN       39     47       EFFECTOR REGION (PROBABLE).
FT   LIPID       200    200       S-farnesyl cysteine (By similarity).
SQ   SEQUENCE   203 AA;  23236 MW;  52098F53F3966A54 CRC64;

  Query Match             65.2%;  Score 43;  DB 1;  Length 203;
  Best Local Similarity   53.3%;  Pred. No. 5.8;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||:| :||:|: :|
Db         17 GGGGVGKSALTIQFI 31


Search completed: June  2, 2004, 18:10:19
Job time : 4.16279 secs
```

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:09:54 ; Search time 13.1783 Seconds
(without alignments)
407.018 Million cell updates/sec

Title: US-10-092-367-6
Perfect score: 66
Sequence: 1 GGGXVRXSAXTLHXITX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	72.7	1417	16	Q9HWR8	Q9hwr8 pseudomonas
2	46	69.7	403	3	O74962	O74962 schizosacch
3	46	69.7	419	2	Q9RNH3	Q9rnh3 rhodobacter
4	46	69.7	722	5	Q9U0Z4	Q9u0z4 leishmania
5	45	68.2	34	13	Q8QGG0	Q8qgg0 oncorhynchu
6	45	68.2	339	16	Q83MJ7	Q83mj7 shigella fl
7	45	68.2	2515	16	Q7UZ67	Q7uz67 rhodopirell
8	44	66.7	113	10	Q7X710	Q7x710 oryza sativ
9	44	66.7	202	12	Q919I7	Q919i7 culex nigri
10	44	66.7	217	16	Q7U203	Q7u203 mycobacteri
11	44	66.7	218	16	P96280	P96280 mycobacteri
12	44	66.7	496	13	Q7SX93	Q7sx93 brachydanio
13	44	66.7	791	3	O43129	O43129 aspergillus
14	44	66.7	792	17	Q9HPN8	Q9hpn8 halobacteri
15	44	66.7	22152	4	Q8WXI7	Q8wxi7 homo sapien
16	43	65.2	168	5	Q8ITX9	Q8itx9 caenorhabdi
17	43	65.2	176	10	Q9XHV9	Q9xhv9 oryza sativ
18	43	65.2	186	5	Q01208	Q01208 dictyosteli
19	43	65.2	191	5	O97342	O97342 suberites d
20	43	65.2	191	5	Q24471	Q24471 dictyosteli
21	43	65.2	203	5	Q24807	Q24807 entamoeba h
22	43	65.2	204	11	Q9D0H6	Q9d0h6 mus musculu
23	43	65.2	204	11	Q8C5D1	Q8c5d1 mus musculu
24	43	65.2	205	5	Q24806	Q24806 entamoeba h
25	43	65.2	210	3	Q9UVQ4	Q9uvq4 cryptococcu
26	43	65.2	210	3	Q9HFU0	Q9hfu0 cryptococcu
27	43	65.2	212	5	O45056	O45056 caenorhabdi
28	43	65.2	213	3	Q9C1I6	Q9c1i6 pisolithus
29	43	65.2	215	3	Q875L4	Q875l4 ustilago ma
30	43	65.2	216	3	Q9P8I9	Q9p8i9 suillus bov
31	43	65.2	289	3	Q9UVU4	Q9uvu4 candida alb
32	43	65.2	309	5	Q9N9D3	Q9n9d3 physarum po
33	43	65.2	310	5	Q9N9D2	Q9n9d2 physarum po
34	43	65.2	343	16	Q8A7X4	Q8a7x4 bacteroides
35	43	65.2	439	4	Q9H5R8	Q9h5r8 homo sapien
36	43	65.2	551	16	Q8FLU0	Q8flu0 corynebacte
37	43	65.2	571	6	Q9BE48	Q9be48 macaca fasc
38	43	65.2	785	4	Q86W01	Q86w01 homo sapien
39	43	65.2	1076	4	Q8N3A9	Q8n3a9 homo sapien
40	43	65.2	1220	11	Q8CDG3	Q8cdg3 mus musculu
41	43	65.2	1221	11	Q8CF97	Q8cf97 rattus norv
42	43	65.2	1221	11	Q7TNK5	Q7tnk5 rattus norv
43	43	65.2	1222	4	Q86T93	Q86t93 homo sapien
44	43	65.2	1236	4	Q96JH7	Q96jh7 homo sapien
45	43	65.2	1245	11	Q80T83	Q80t83 mus musculu

ALIGNMENTS

RESULT 1
Q9HWR8
ID Q9HWR8 PRELIMINARY; PRT; 1417 AA.
AC Q9HWR8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable sensor/response regulator hybrid.
GN PA4112.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL; AE004827; AAG07499.1; -.
DR PIR; H83132; H83132.
DR HSSP; P06143; 1AB6.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.

```
DR   GO; GO:0007600; P:sensory perception; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR006189; CHASE.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR003661; His_kinA_N.
DR   InterPro; IPR008207; Hpt.
DR   InterPro; IPR001610; PAC.
DR   InterPro; IPR000700; PAS-assoc_C.
DR   InterPro; IPR000014; PAS_domain.
DR   InterPro; IPR001789; Response_reg.
DR   Pfam; PF03924; CHASE; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00512; HisKA; 1.
DR   Pfam; PF01627; Hpt; 1.
DR   Pfam; PF00785; PAC; 3.
DR   Pfam; PF00989; PAS; 2.
DR   Pfam; PF00072; response_reg; 2.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD000039; Response_reg; 2.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00073; HPT; 1.
DR   SMART; SM00086; PAC; 3.
DR   SMART; SM00091; PAS; 3.
DR   SMART; SM00448; REC; 2.
DR   TIGRFAMs; TIGR00229; sensory_box; 3.
DR   PROSITE; PS50839; CHASE; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50894; HPT; 1.
DR   PROSITE; PS50113; PAC; 3.
DR   PROSITE; PS50112; PAS; 2.
DR   PROSITE; PS50110; RESPONSE_REGULATORY; 2.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase;
KW   Complete proteome.
SQ   SEQUENCE   1417 AA;  153893 MW;  224E2EC9E45EAF2B CRC64;

  Query Match             72.7%;  Score 48;  DB 16;  Length 1417;
  Best Local Similarity   60.0%;  Pred. No. 1e+02;
  Matches    9;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              | |:|::||:|||:|
Db       1336 GEGDVQGSAATLHTI 1350


RESULT 2
O74962
ID   O74962      PRELIMINARY;      PRT;   403 AA.
AC   O74962;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative pantothenate kinase.
GN   SPBC4B4.01C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972h-;
RA   Beck A., Reinhardt R, Lyne M., Wood V., Rajandream M.A., Barrell B.G.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL023706; CAA19281.1; -.
DR   PIR; T40473; T40473.
DR   GeneDB_SPombe; SPBC4B4.01c; -.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   InterPro; IPR004567; PanK_eukar.
DR   Pfam; PF03630; Fumble; 1.
DR   TIGRFAMs; TIGR00555; panK_eukar; 1.
KW   Kinase.
SQ   SEQUENCE   403 AA;  44861 MW;  E4574392867BFE20 CRC64;

  Query Match             69.7%;  Score 46;  DB 3;  Length 403;
  Best Local Similarity   35.3%;  Pred. No. 54;
  Matches    6;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              || ::|:  :|:|::|:
Db        343 GGSFIRNHVQTMHTLTY 359


RESULT 3
Q9RNH3
ID   Q9RNH3      PRELIMINARY;      PRT;   419 AA.
AC   Q9RNH3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Sensor kinase homolog (Fragment).
GN   CCKA.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B10;
RX   MEDLINE=20105563; PubMed=10639170;
RA   Lang A.S., Beatty J.T.;
RT   "Genetic analysis of a bacterial genetic exchange element: The gene
RT   transfer agent of Rhodobacter capsulatus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:859-864(2000).
CC   -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC       KINASES.
DR   EMBL; AF181079; AAF13178.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0000156; F:two-component response regulator activity; IEA.
DR   GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR   GO; GO:0007600; P:sensory perception; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR003661; His_kinA_N.
DR   InterPro; IPR001789; Response_reg.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00512; HisKA; 1.
DR   Pfam; PF00072; response_reg; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD000039; Response_reg; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00448; REC; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase.
FT   NON_TER       1      1
SQ   SEQUENCE   419 AA;  45836 MW;  9A94A5EF348A39FC CRC64;

  Query Match             69.7%;  Score 46;  DB 2;  Length 419;
  Best Local Similarity   46.7%;  Pred. No. 57;
  Matches    7;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||::|:  : ||:|
Db        170 GGGEIRIETENLHLI 184
```

```
RESULT 4
Q9U0Z4
ID   Q9U0Z4      PRELIMINARY;      PRT;   722 AA.
AC   Q9U0Z4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   L5883.03.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RA   Murphy L., Harris D., Ivens A.C., Lawson D., Quail M.,
RA   Rajandream M.A., Barrell B.G.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL117384; CAB55614.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   InterPro; IPR000330; SNF2_N.
DR   Pfam; PF00176; SNF2_N; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   722 AA;  74613 MW;  1AFEDBBF764DF361 CRC64;

  Query Match             69.7%;  Score 46;  DB 5;  Length 722;
  Best Local Similarity   46.7%;  Pred. No. 1e+02;
  Matches    7;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              |||: |:||:::|::
Db        227 GGGAPRASANSVHGV 241


RESULT 5
Q8QGG0
ID   Q8QGG0      PRELIMINARY;      PRT;    34 AA.
AC   Q8QGG0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   K-ras (Fragment).
OS   Oncorhynchus gorbuscha (Pink salmon) (Humpback salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8017;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PWS8;
RA   Cronin M.A., Wickliffe J.K., Dunina Y., Baker R.J.;
RT   "K-ras oncogene DNA sequences in pink salmon in streams impacted by
RT   the Exxon Valdez oil spill: no evidence of oil-induced heritable
RT   mutations.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF465435; AAM11562.1; -.
DR   GO; GO:0005525; F:GTP binding; IEA.
DR   GO; GO:0003925; F:small monomeric GTPase activity; IEA.
DR   GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
KW   GTP-binding.
FT   NON_TER      34      34
SQ   SEQUENCE   34 AA;  3702 MW;  41D73D6875AE4F4F CRC64;

  Query Match             68.2%;  Score 45;  DB 13;  Length 34;
  Best Local Similarity   53.3%;  Pred. No. 5.3;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXI 15
              | |:| :||:|:|:|
Db         10 GAGGVGKSALTIHLI 24


RESULT 6
Q83MJ7
ID   Q83MJ7      PRELIMINARY;      PRT;   339 AA.
AC   Q83MJ7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chain B, crystal structure of Rna 3'-terminal phosphate cyclase, An
DE   ubiquitous enzyme with Unusual Topology.
GN   SF3442.
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
DR   EMBL; AE015352; AAN44902.1; -.
DR   GO; GO:0003963; F:RNA-3'-phosphate cyclase activity; IEA.
DR   InterPro; IPR000228; RNA3'_term_cycl.
DR   Pfam; PF01137; RTC; 1.
DR   Pfam; PF05189; RTC_insert; 1.
DR   PROSITE; PS01287; RTC; 1.
KW   Complete proteome.
SQ   SEQUENCE   339 AA;  36024 MW;  AD7F57EF111CD266 CRC64;

  Query Match             68.2%;  Score 45;  DB 16;  Length 339;
  Best Local Similarity   47.1%;  Pred. No. 66;
  Matches    8;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:: :||::| :||:
Db         15 GGGQILRSALSLSMITG 31


RESULT 7
Q7UZ67
ID   Q7UZ67      PRELIMINARY;      PRT;  2515 AA.
AC   Q7UZ67;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Lipoprotein receptor-related protein.
GN   RB154.
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
OX   NCBI_TaxID=117;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
```

```
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294133; CAD71416.1; -.
KW   Lipoprotein; Receptor; Complete proteome.
SQ   SEQUENCE   2515 AA;  261824 MW;  C319023DC36D6762 CRC64;

  Query Match             68.2%;  Score 45;  DB 16;  Length 2515;
  Best Local Similarity   41.2%;  Pred. No. 6e+02;
  Matches    7;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:  : |:||:::|:
Db        464 GGGITNRGAATLNRVTL 480


RESULT 8
Q7X710
ID   Q7X710      PRELIMINARY;      PRT;   113 AA.
AC   Q7X710;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   OSJNBa0016N04.23 protein (OSJNBb0042I07.8 protein).
GN   OSJNBA0016N04.23 OR OSJNBB0042I07.8.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA   Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA   Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA   Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA   Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA   Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA   Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA   Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA   Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL731587; CAD40621.1; -.
DR   EMBL; AL731632; CAD40711.1; -.
SQ   SEQUENCE   113 AA;  12012 MW;  A54BBBB801B51727 CRC64;

  Query Match             66.7%;  Score 44;  DB 10;  Length 113;
  Best Local Similarity   52.9%;  Pred. No. 29;
  Matches    9;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:||:||:|  : |:
Db         70 GGGSVRCSATTRTVSTF 86


RESULT 9
Q919I7
ID   Q919I7      PRELIMINARY;      PRT;   202 AA.
AC   Q919I7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CUN090 putative similar to AcMNPV ORF96.
GN   CUN090.
OS   Culex nigripalpus baculovirus.
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae.
OX   NCBI_TaxID=130556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Florida1997;
RX   MEDLINE=21488685; PubMed=11602755;
RA   Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
RA   Becnel J.J., Rock D.L., Kutish G.F.;
RT   "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus.";
RL   J. Virol. 75:11157-11165(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Florida1997;
RA   Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
RA   Becnel J.J., Rock D.L., Kutish G.F.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF403738; AAK94168.1; -.
DR   InterPro; IPR006883; Baculo_19.
DR   Pfam; PF04798; Baculo_19; 1.
SQ   SEQUENCE   202 AA;  23082 MW;  115F79E4BF667E88 CRC64;

  Query Match             66.7%;  Score 44;  DB 12;  Length 202;
  Best Local Similarity   66.7%;  Pred. No. 55;
  Matches    8;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTL 12
              |||:||::|:||
Db         30 GGGIVRHAADTL 41


RESULT 10
Q7U203
ID   Q7U203      PRELIMINARY;      PRT;   217 AA.
AC   Q7U203;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Conserved hypothetical protein.
GN   MB0442.
OS   Mycobacterium bovis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1765;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR   EMBL; BX248335; CAD93305.1; -.
KW   Complete proteome.
SQ   SEQUENCE   217 AA;  23720 MW;  40C8B116384F15C7 CRC64;

  Query Match             66.7%;  Score 44;  DB 16;  Length 217;
  Best Local Similarity   47.1%;  Pred. No. 60;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||: |:  :|| :||:
Db         60 GGGDTRCDVGTLARITE 76


RESULT 11
P96280
ID   P96280      PRELIMINARY;      PRT;   218 AA.
AC   P96280;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein Rv0434.
GN   RV0434 OR MT0449 OR MTCY22G10.31.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z84724; CAB06574.1; ALT_INIT.
DR   EMBL; AE006948; AAK44672.1; -.
DR   PIR; H70631; H70631.
DR   TIGR; MT0449; -.
DR   TubercuList; Rv0434; -.
DR   GO; GO:0004176; F:ATP-dependent peptidase activity; IEA.
DR   GO; GO:0006510; P:ATP-dependent proteolysis; IEA.
DR   InterPro; IPR003111; Pept_S16_N.
DR   Pfam; PF02190; LON; 1.
DR   SMART; SM00464; LON; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   218 AA;  23881 MW;  7088BC9B292EE877 CRC64;

  Query Match             66.7%;  Score 44;  DB 16;  Length 218;
  Best Local Similarity   47.1%;  Pred. No. 60;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||: |:  :|| :||:
Db         61 GGGDTRCDVGTLARITE 77


RESULT 12
Q7SX93
ID   Q7SX93      PRELIMINARY;      PRT;   496 AA.
AC   Q7SX93;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC055582; AAH55582.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   496 AA;  55845 MW;  93329B64B2A448A0 CRC64;

  Query Match             66.7%;  Score 44;  DB 13;  Length 496;
  Best Local Similarity   41.2%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              || ::|:  :|:|:||:
Db        322 GGFFIRGHPVTMHTITY 338


RESULT 13
O43129
ID   O43129      PRELIMINARY;      PRT;   791 AA.
AC   O43129;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Multidrug resistance protein 2.
GN   MDR2.
OS   Aspergillus fumigatus (Sartorya fumigata).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5085;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=10AF/86/10;
RX   MEDLINE=98038972; PubMed=9373135;
RA   Tobin M.B., Peery R.B., Skatrud P.L.;
RT   "Genes encoding multiple drug resistance-like proteins in Aspergillus
RT   fumigatus and Aspergillus flavus.";
RL   Gene 200:11-23(1997).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR   EMBL; U62936; AAB88660.1; -.
DR   EMBL; U62935; AAB88659.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001140; ABC_TM_transpt.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   ATP-binding; Transport.
SQ   SEQUENCE   791 AA;  85195 MW;  EAA0E5535CC3BCF0 CRC64;

  Query Match             66.7%;  Score 44;  DB 3;  Length 791;
  Best Local Similarity   41.2%;  Pred. No. 2.5e+02;
```

```
  Matches    7;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:|:: |:|: ::|:
Db        431 GGGMVQSGAITIGELTS 447


RESULT 14
Q9HPN8
ID   Q9HPN8      PRELIMINARY;      PRT;   792 AA.
AC   Q9HPN8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chloride channel.
GN   CLC OR VNG1544G.
OS   Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC   Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC   Halobacteriaceae; Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005067; AAG19829.1; -.
DR   PIR; A84308; A84308.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005247; F:voltage-gated chloride channel activity; IEA.
DR   GO; GO:0006821; P:chloride transport; IEA.
DR   InterPro; IPR000644; CBS_domain.
DR   InterPro; IPR001807; Cl-channel_volt.
DR   Pfam; PF00571; CBS; 2.
DR   Pfam; PF00654; voltage_CLC; 1.
DR   PRINTS; PR00762; CLCHANNEL.
DR   SMART; SM00116; CBS; 2.
KW   Complete proteome.
SQ   SEQUENCE   792 AA;  81682 MW;  77E1D7CB2635CD46 CRC64;

  Query Match             66.7%;  Score 44;  DB 17;  Length 792;
  Best Local Similarity   47.1%;  Pred. No. 2.5e+02;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              |||:  :||: ||:| :
Db         57 GGGLAVVSAYNLHRIAH 73


RESULT 15
Q8WXI7
ID   Q8WXI7      PRELIMINARY;      PRT; 22152 AA.
AC   Q8WXI7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ovarian cancer related tumor marker CA125.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21646939; PubMed=11786729;
RA   O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
RA   York L.;
RT   "The CA125 gene: An extracellular superstructure dominated by repeat
RT   sequences.";
RL   Tumour Biol. 22:348-366(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   O'Brien T.J., Underwood L.J., Beard J.B.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF414442; AAL65133.2; -.
DR   Genew; HGNC:15582; MUC16.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0015992; P:proton transport; IEA.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR000082; SEA_domain.
DR   Pfam; PF01390; SEA; 51.
DR   SMART; SM00200; SEA; 23.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 2.
DR   PROSITE; PS50024; SEA; 11.
SQ   SEQUENCE   22152 AA;  2352797 MW;  B3E7BDF19997A440 CRC64;

  Query Match             66.7%;  Score 44;  DB 4;  Length 22152;
  Best Local Similarity   41.2%;  Pred. No. 9.6e+03;
  Matches    7;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGXVRXSAXTLHXITX 17
              || : |:| :| |::|:
Db        613 GGSSTRGSQGTTHLLTR 629


Search completed: June  2, 2004, 18:12:13
Job time : 15.1783 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 17:58:08 ; Search time 106.783 Seconds
(without alignments)
251.370 Million cell updates/sec

Title: US-10-092-367-73
Perfect score: 479
Sequence: 1 MQLYTYLYLLVPLVTFYLIL.....GNMRGGEVRESAETLHEITP 95

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	479	100.0	95	6	ABJ38902	Abj38902 Conopepti
2	471	98.3	95	6	ABJ38896	Abj38896 Conopepti
3	469	97.9	95	6	ABJ38894	Abj38894 Conopepti
4	436	91.0	97	6	ABJ38898	Abj38898 Conopepti
5	405.5	84.7	94	6	ABJ38922	Abj38922 Conopepti
6	385	80.4	100	6	ABJ38914	Abj38914 Conopepti
7	379	79.1	98	6	ABJ38900	Abj38900 Conopepti
8	377	78.7	101	4	AAU01508	Aau01508 Propeptid
9	371	77.5	101	6	ABJ38924	Abj38924 Conopepti
10	364.5	76.1	102	6	ABJ38880	Abj38880 Conopepti
11	363	75.8	107	6	ABJ38888	Abj38888 Conopepti
12	362.5	75.7	102	4	AAU01510	Aau01510 Propeptid
13	362.5	75.7	107	2	AAW48211	Aaw48211 Conus rad
14	362.5	75.7	107	2	AAW49990	Aaw49990 Conus rad
15	362.5	75.7	107	4	AAG79045	Aag79045 Amino aci
16	361.5	75.5	102	4	AAU01509	Aau01509 Propeptid
17	359	74.9	96	6	ABJ38906	Abj38906 Conopepti
18	359	74.9	99	6	ABJ38920	Abj38920 Conopepti
19	358.5	74.8	102	6	ABJ38878	Abj38878 Conopepti
20	356	74.3	99	6	ABJ38908	Abj38908 Conopepti
21	354	73.9	96	6	ABJ38938	Abj38938 Conopepti
22	353	73.7	100	2	AAW48210	Aaw48210 Conus geo
23	353	73.7	100	2	AAW49989	Aaw49989 Conus geo
24	353	73.7	100	2	AAY30355	Aay30355 A conanto
25	353	73.7	100	4	AAU01503	Aau01503 Propeptid
26	353	73.7	100	4	AAG79044	Aag79044 Amino aci
27	352.5	73.6	102	6	ABJ38882	Abj38882 Conopepti
28	352	73.5	99	6	ABJ38912	Abj38912 Conopepti
29	350.5	73.2	101	6	ABJ38890	Abj38890 Conopepti
30	350	73.1	102	6	ABJ38930	Abj38930 Conopepti
31	349	72.9	98	4	AAU01516	Aau01516 Propeptid
32	348	72.7	102	6	ABJ38934	Abj38934 Conopepti
33	347.5	72.5	103	6	ABJ38892	Abj38892 Conopepti
34	347.5	72.5	114	6	ABJ38918	Abj38918 Conopepti
35	347.5	72.5	114	6	ABJ38916	Abj38916 Conopepti
36	347	72.4	103	2	AAW48213	Aaw48213 Conus sul
37	347	72.4	103	2	AAW49992	Aaw49992 Conus sul
38	347	72.4	103	4	AAG79047	Aag79047 Amino aci
39	345	72.0	96	4	AAU01517	Aau01517 Propeptid
40	343	71.6	101	6	ABJ38910	Abj38910 Conopepti
41	343	71.6	111	4	AAU01702	Aau01702 Propeptid
42	343	71.6	111	4	AAU01701	Aau01701 Propeptid
43	342	71.4	98	4	AAU01518	Aau01518 Propeptid
44	341	71.2	100	4	AAU01507	Aau01507 Propeptid
45	340.5	71.1	97	6	ABJ38940	Abj38940 Conopepti

ALIGNMENTS

RESULT 1
ABJ38902
ID ABJ38902 standard; protein; 95 AA.
XX
AC ABJ38902;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Bt5 SEQ ID No 73.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43476.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX

PS Claim 5; Page 33; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 95 AA;

```
  Query Match             100.0%;  Score 479;  DB 6;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-54;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
              |||||||||||||||||||||||||||||||||||
Db         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
```

RESULT 2
ABJ38896
ID ABJ38896 standard; protein; 95 AA.
XX
AC ABJ38896;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Bt2 SEQ ID No 64.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43473.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 32; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 95 AA;

```
  Query Match             98.3%;  Score 471;  DB 6;  Length 95;
  Best Local Similarity   98.9%;  Pred. No. 4.7e-53;
  Matches   94;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
```

```
Db         61 GKDRLTQMIRILKKRGNMRGEEVRESAETLHEITP 95

RESULT 3
ABJ38894
ID   ABJ38894 standard; protein; 95 AA.
XX
AC   ABJ38894;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide conotoxin protein Bt1 SEQ ID No 61.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
DR   N-PSDB; ABT43472.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Claim 5; Page 31; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin protein of the invention
XX
SQ   Sequence 95 AA;

  Query Match             97.9%;  Score 469;  DB 6;  Length 95;
  Best Local Similarity   97.9%;  Pred. No. 8.5e-53;
  Matches   93;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
              |||||||||||||||||||| |||||||||||:||
Db         61 GKDRLTQMIRILKKRGNMRGEEVRESAETLHELTP 95


RESULT 4
ABJ38898
ID   ABJ38898 standard; protein; 97 AA.
XX
AC   ABJ38898;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide conotoxin protein Bt3 SEQ ID No 67.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW   parasitic worm.
XX
OS   Conus betulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
DR   N-PSDB; ABT43474.
XX
```

PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 32; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 97 AA;

Query Match 91.0%; Score 436; DB 6; Length 97;
Best Local Similarity 90.7%; Pred. No. 1.7e-48;
Matches 88; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

Qy 1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
Db 1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPKPILLQKSAARSTDDN 60

Qy 61 GKDRLTQMIRILKKRGNM--RGGEVRESAETLHEITP 95
Db 61 GKDRLTQMIRILKKRGNMGRDGEEVREAAETLNELTP 97

RESULT 5
ABJ38922
ID ABJ38922 standard; protein; 94 AA.
XX
AC ABJ38922;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Fi4 SEQ ID No 103.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus figulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43486.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 38; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 94 AA;

Query Match 84.7%; Score 405.5; DB 6; Length 94;
Best Local Similarity 88.3%; Pred. No. 1.5e-44;
Matches 83; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEIT 94
              ||||||||   :|||||    ||||:||||||::
Db         61 GKDRLTQMKGTVKKRGN-TAEEVREAAETLHELS 93
```

RESULT 6
ABJ38914
ID ABJ38914 standard; protein; 100 AA.
XX
AC ABJ38914;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Fi1 SEQ ID No 91.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus figulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43482.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 36; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 100 AA;

Query Match 80.4%; Score 385; DB 6; Length 100;
Best Local Similarity 90.8%; Pred. No. 7.6e-42;
Matches 79; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESA 87
              ||||||| || ||||| |  || | |
Db         61 DKDRLTQMKRIFKKRGNKREEEVAEMA 87
```

RESULT 7
ABJ38900
ID ABJ38900 standard; protein; 98 AA.
XX
AC ABJ38900;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein Bt4 SEQ ID No 70.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus betulinus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;

PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43475.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 33; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 98 AA;

```
  Query Match             79.1%;  Score 379;  DB 6;  Length 98;
  Best Local Similarity   83.9%;  Pred. No. 4.5e-41;
  Matches   78;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:|||||||||||||||| | |||||||| ||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLGHGGALTESRSADATALKPGPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEI 93
              |||||||| | |||||| |  : || |||: |:
Db         61 GKDRLTQMKRTLKKRGNTRYEDDREIAETVREL 93
```

RESULT 8
AAU01508
ID AAU01508 standard; protein; 101 AA.
XX
AC AAU01508;
XX
DT 29-AUG-2001 (first entry)
XX
DE Propeptide of conopeptide O2, amino acid sequence.
XX
KW Gamma carboxyglutamate; neurological disorder; epilepsy; trauma; hypoxia;
KW anoxia; ischaemia; stroke; brain; spinal cord; suffocation;
KW myocardial infarct; drowning; perinatal asphyxia; hypoglycaemia;
KW neurodegeneration; Alzheimer's disease; Huntington's disease;
KW senile dementia; Amyotrophic Lateral Sclerosis; multiple sclerosis;
KW Parkinson's disease; Down's Syndrome; Korsakoff's disease; schizophrenia;
KW AIDS; acquired immunodeficiency syndrome; HIV; neuronal damage; pain;
KW seizure; chemical toxicity; addiction; dystonia; psychiatric disorder;
KW mood disorder; memory; ophthalmic; parasitic worm; conopeptide O2.
XX
OS Conus obscurus.
XX
PN WO200118033-A1.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000WO-US024816.
XX
PR 10-SEP-1999; 99US-0153034P.
PR 21-JUL-2000; 2000US-0219673P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2001-273379/28.
DR N-PSDB; AAS02195.
XX
PT New isolated gamma-carboxyglutamine containing peptide for treating or
PT preventing neurological and psychiatric disorders e.g. epilepsy,
PT Alzheimer's disease, migraine, chemical toxicity, dystonia, anxiety, and
PT depression.
XX
PS Claim 5; Page 33; 102pp; English.
XX
CC The sequence represents the amino acid sequence of the propeptide of
CC gamma carboxyglutamate-containing conopeptide O2. The conopeptide is used
CC for treating or preventing disorders in which the pathophysiology
CC involves excess excitation of nerve cells by excitatory amino acids or
CC agonists of heterogenous ionotropic glutamate receptors or heterogenous G
CC protein coupled glutamate receptors. The disorders may be neurological
CC disorders, such as: (i) seizure associated with epilepsy; (ii) a
CC neurotoxic injury associated with hypoxia, anoxia, ischaemia, stroke,
CC cerebrovascular accident, brain or spinal cord trauma, myocardial
CC infarct, physical trauma, drownings, suffocation, perinatal asphyxia, or
CC hypoglycaemic events; (iii) neurodegeneration associated with Alzheimer's
CC disease, Huntington's disease, senile dementia, Amyotrophic Lateral
CC Sclerosis, multiple sclerosis, Parkinson's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS (acquired immunodeficiency
CC syndrome) dementia from HIV infection, HIV infection, multi-infarct
CC dementia, Binswanger dementia and neuronal damage associated with
CC uncontrolled seizures; (iv) pain which is a migraine, acute pain, or
CC persistent pain; (v) chemical toxicity which is addiction, morphine,
CC opiate, opioid and barbiturate tolerance; and (vi) dystonia, urinary
CC incontinence, muscle relaxation or sleep disorder. The disorders may be
CC psychiatric disorders, such as, anxiety, major depression, manic-
CC depressive illness, obsessive-compulsive disorder, schizophrenia, or mood
CC disorders (bipolar disorder, unipolar depression, dysthymia, or seasonal
CC effective disorder). The conopeptide is also used to treat memory or
CC cognitive deficits, ophthalmic indications, or to control nematodes or
CC parasitic worms
XX
SQ Sequence 101 AA;

```
  Query Match             78.7%;  Score 377;  DB 4;  Length 101;
  Best Local Similarity   83.2%;  Pred. No. 8.5e-41;
  Matches   79;  Conservative    4;  Mismatches   10;  Indels    2;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:|||||||| |||||||| ||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGN--MRGGEVRESAETLHEI 93
```

```
                ||||||||  |||||||    |  |:||  ||: |:
Db         61 GKDRLTQMKGILKKRGNTARRDEELREDVETILEL 95

RESULT 9
ABJ38924
ID   ABJ38924 standard; protein; 101 AA.
XX
AC   ABJ38924;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide conotoxin protein Fi5 SEQ ID No 106.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW   parasitic worm.
XX
OS   Conus figulinus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
DR   N-PSDB; ABT43487.
XX
PT   New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT   seizure associated with epilepsy or neurotoxic injury associated with
PT   hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT   morphine tolerance).
XX
PS   Claim 5; Page 38; 113pp; English.
XX
CC   This invention relates to a novel isolated peptide consisting of
CC   conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC   Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC   Sm1. The isolated conotoxin peptides are useful in methods for treating
CC   or preventing disorders in which the pathophysiology involves excessive
CC   excitation of nerve cells by excitatory amino acids or agonists of
CC   heterogenous inotropic glutamate receptors or heterogenous B protein
CC   coupled glutamate receptors; and for treating memory or cognitive
CC   deficits, HIV infection, or ophthalmic indications comprising
CC   administering to a patient a peptide above or its salt. Disorders include
CC   neurological disorder or a psychiatric disorder, where the neurological
CC   disorder is seizure associated with epilepsy or neurotoxic injury
CC   associated with conditions of hypoxia, anoxia or ischaemia, including
CC   neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC   or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC   suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC   disorder may also be a neurodegeneration associated with Alzheimer's
CC   disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC   Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC   Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC   multi-infarct dementia, Binswanger dementia and neuronal damage
CC   associated with uncontrolled seizures. The neurologic disorder is pain
CC   (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC   addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC   barbiturate tolerance), dystonia (movement disorder), urinary
CC   incontinence, muscle relaxation or sleep disorder. The psychiatric
CC   disorder is anxiety, major depression, manic-depressive illness,
CC   obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC   bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC   disorder. The conotoxin peptides are also useful for controlling
CC   nematodes or parasitic worms by applying the peptides to the locus to be
CC   protected. This sequence represents a linear gamma-carboxyglutamate rich
CC   conotoxin protein of the invention
XX
SQ   Sequence 101 AA;

  Query Match             77.5%;  Score 371;  DB 6;  Length 101;
  Best Local Similarity   82.1%;  Pred. No. 5.1e-40;
  Matches   78;  Conservative    6;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||| |
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDVN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHEI 93
              ||||||:| |||||||::  |    || || : |:
Db         61 GKDRLTEMKRILKKRGSISMGFEHRREIAELVREL 95


RESULT 10
ABJ38880
ID   ABJ38880 standard; protein; 102 AA.
XX
AC   ABJ38880;
XX
DT   09-OCT-2003  (first entry)
XX
DE   Conopeptide conotoxin protein C-2 SEQ ID No 40.
XX
KW   Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW   antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW   tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW   Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW   Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW   inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW   heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW   seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW   neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW   myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW   hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW   pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW   parasitic worm.
XX
OS   Conus catus.
XX
PN   WO200272005-A2.
XX
PD   19-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-US006863.
XX
PR   07-MAR-2001; 2001US-0273639P.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Mcintosh JM,  Garrett JE,  Walker CS,  Watkins M;
PI   Jones RM;
XX
DR   WPI; 2003-175000/17.
DR   N-PSDB; ABT43465.
XX
```

PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 28; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 102 AA;

```
  Query Match             76.1%;  Score 364.5;  DB 6;  Length 102;
  Best Local Similarity   78.1%;  Pred. No. 3.6e-39;
  Matches   75;  Conservative    8;  Mismatches   10;  Indels    3;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||:||||||| |||||||||  |||||:||||||||||||||||:
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSGDATALRPEPVLLQKSAARSTDDS 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHEI 93
              |||||||| |||||:||   |:   :||  ||: |:
Db         61 GKDRLTQMKRILKKQGNTAKGDEELLREDVETVLEL 96
```

RESULT 11
ABJ38888
ID ABJ38888 standard; protein; 107 AA.
XX
AC ABJ38888;
XX
DT 09-OCT-2003 (first entry)
XX
DE Conopeptide conotoxin protein C-6 SEQ ID No 52.
XX
KW Neuroprotective; anticonvulsant; cerebroprotective; cardiant; analgesic;
KW antidiabetic; nootropic; anti-Parkinsonian; antiaddictive; vasotropic;
KW tranquiliser; antidepressant; peptide therapy; conotoxin; Af6; Bt1; Bt2;
KW Bt3; Bt4; Bt5; Bu1; Bu2; C1; C2; C3; C4; C5; C6; Di1; Di2; Ep1; Fi1; Fi2;
KW Fi3; Fi4; Fi5; L1; L2; L3; P1; P2; P3; P4; P5; Sm1; nerve cell; memory;
KW inotropic glutamate receptor; neurological disorder; cognitive; deficit;
KW heterogenous B protein coupled glutamate receptor; HIV; psychiatric;
KW seizure; epilepsy; neurotoxic injury; hypoxia; anoxia; ischaemia; stroke;
KW neurotoxic injury; cerebrovascular accident; brain; spinal cord trauma;
KW myocardial infarct; physical trauma; drowning; suffocation; dystonia;
KW hypoglycaemic; perinatal asphyxia; neurodegeneration; chemical toxicity;
KW pain; nematode; linear gamma-carboxyglutamate rich conotoxin;
KW parasitic worm.
XX
OS Conus catus.
XX
PN WO200272005-A2.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US006863.
XX
PR 07-MAR-2001; 2001US-0273639P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2003-175000/17.
DR N-PSDB; ABT43469.
XX
PT New conotoxins useful for treating e.g. neurologic disorders (e.g.
PT seizure associated with epilepsy or neurotoxic injury associated with
PT hypoxia), pain (e.g. migraine), or chemical toxicity (e.g. addiction or
PT morphine tolerance).
XX
PS Claim 5; Page 30; 113pp; English.
XX
CC This invention relates to a novel isolated peptide consisting of
CC conotoxin Af6, Bt1, Bt2, Bt3, Bt4, Bt5, Bu1, Bu2, C1, C2, C3, C4, C5, C6,
CC Di1, Di2, Ep1, Fi1, Fi2, Fi3, Fi4, Fi5, L1, L2, L3, P1, P2, P3, P4, P5 or
CC Sm1. The isolated conotoxin peptides are useful in methods for treating
CC or preventing disorders in which the pathophysiology involves excessive
CC excitation of nerve cells by excitatory amino acids or agonists of
CC heterogenous inotropic glutamate receptors or heterogenous B protein
CC coupled glutamate receptors; and for treating memory or cognitive
CC deficits, HIV infection, or ophthalmic indications comprising
CC administering to a patient a peptide above or its salt. Disorders include
CC neurological disorder or a psychiatric disorder, where the neurological
CC disorder is seizure associated with epilepsy or neurotoxic injury
CC associated with conditions of hypoxia, anoxia or ischaemia, including
CC neurotoxic injury associated with stroke, cerebrovascular accident, brain
CC or spinal cord trauma, myocardial infarct, physical trauma, drownings,
CC suffocation, perinatal asphyxia or hypoglycaemic events. The neurologic
CC disorder may also be a neurodegeneration associated with Alzheimer's
CC disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple
CC Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS dementia from HIV infection,
CC multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures. The neurologic disorder is pain
CC (e.g. migraine, acute pain or persistent pain), chemical toxicity (e.g.
CC addiction, morphine tolerance, opiate tolerance, opioid tolerance and
CC barbiturate tolerance), dystonia (movement disorder), urinary
CC incontinence, muscle relaxation or sleep disorder. The psychiatric
CC disorder is anxiety, major depression, manic-depressive illness,
CC obsessive-compulsive disorder, schizophrenia or a mood disorder, such as
CC bipolar disorder, unipolar depression, dysthymia or seasonal effective
CC disorder. The conotoxin peptides are also useful for controlling
CC nematodes or parasitic worms by applying the peptides to the locus to be
CC protected. This sequence represents a linear gamma-carboxyglutamate rich
CC conotoxin protein of the invention
XX
SQ Sequence 107 AA;

```
  Query Match             75.8%;  Score 363;  DB 6;  Length 107;
  Best Local Similarity   81.5%;  Pred. No. 6.1e-39;
  Matches   75;  Conservative    3;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||:|||||||| |||||||| |||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHE 92
              |||||| | |||||| | |  ||    | : :
Db         61 GKDRLTHMKRILKKRANKREPEVGSIPEAVRQ 92
```

RESULT 12
AAU01510
ID AAU01510 standard; protein; 102 AA.
XX
AC AAU01510;
XX
DT 29-AUG-2001 (first entry)
XX
DE Propeptide of conopeptide O2B, amino acid sequence.
XX
KW Gamma carboxyglutamate; neurological disorder; epilepsy; trauma; hypoxia;
KW anoxia; ischaemia; stroke; brain; spinal cord; suffocation;
KW myocardial infarct; drowning; perinatal asphyxia; hypoglycaemia;
KW neurodegeneration; Alzheimer's disease; Huntington's disease;
KW senile dementia; Amyotrophic Lateral Sclerosis; multiple sclerosis;
KW Parkinson's disease; Down's Syndrome; Korsakoff's disease; schizophrenia;
KW AIDS; acquired immunodeficiency syndrome; HIV; neuronal damage; pain;
KW seizure; chemical toxicity; addiction; dystonia; psychiatric disorder;
KW mood disorder; memory; ophthalmic; parasitic worm; conopeptide O2B.
XX
OS Conus obscurus.
XX
PN WO200118033-A1.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000WO-US024816.
XX
PR 10-SEP-1999; 99US-0153034P.
PR 21-JUL-2000; 2000US-0219673P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Garrett JE, Walker CS, Watkins M;
PI Jones RM;
XX
DR WPI; 2001-273379/28.
DR N-PSDB; AAS02197.
XX
PT New isolated gamma-carboxyglutamine containing peptide for treating or
PT preventing neurological and psychiatric disorders e.g. epilepsy,
PT Alzheimer's disease, migraine, chemical toxicity, dystonia, anxiety, and
PT depression.
XX
PS Claim 5; Page 34; 102pp; English.
XX
CC The sequence represents the amino acid sequence of the propeptide of
CC gamma carboxyglutamate-containing conopeptide O2B. The conopeptide is
CC used for treating or preventing disorders in which the pathophysiology
CC involves excess excitation of nerve cells by excitatory amino acids or
CC agonists of heterogenous ionotropic glutamate receptors or heterogenous G
CC protein coupled glutamate receptors. The disorders may be neurological
CC disorders, such as: (i) seizure associated with epilepsy; (ii) a
CC neurotoxic injury associated with hypoxia, anoxia, ischaemia, stroke,
CC cerebrovascular accident, brain or spinal cord trauma, myocardial
CC infarct, physical trauma, drownings, suffocation, perinatal asphyxia, or
CC hypoglycaemic events; (iii) neurodegeneration associated with Alzheimer's
CC disease, Huntington's disease, senile dementia, Amyotrophic Lateral
CC Sclerosis, multiple sclerosis, Parkinson's disease, Down's Syndrome,
CC Korsakoff's disease, schizophrenia, AIDS (acquired immunodeficiency
CC syndrome) dementia from HIV infection, HIV infection, multi-infarct
CC dementia, Binswanger dementia and neuronal damage associated with
CC uncontrolled seizures; (iv) pain which is a migraine, acute pain, or
CC persistent pain; (v) chemical toxicity which is addiction, morphine,
CC opiate, opioid and barbiturate tolerance; and (vi) dystonia, urinary
CC incontinence, muscle relaxation or sleep disorder. The disorders may be
CC psychiatric disorders, such as, anxiety, major depression, manic-
CC depressive illness, obsessive-compulsive disorder, schizophrenia, or mood
CC disorders (bipolar disorder, unipolar depression, dysthymia, or seasonal
CC effective disorder). The conopeptide is also used to treat memory or
CC cognitive deficits, ophthalmic indications, or to control nematodes or
CC parasitic worms
XX
SQ Sequence 102 AA;

Query Match 75.7%; Score 362.5; DB 4; Length 102;
Best Local Similarity 81.2%; Pred. No. 6.6e-39;
Matches 78; Conservative 5; Mismatches 10; Indels 3; Gaps 2;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||:|||||||| |||||||| |||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGN--MRGGE-VRESAETLHEI 93
              |||||||  ||||:||   |  | :||  ||: |:
Db         61 GKDRLTQMKGILKKQGNTARRDEELLREDVETILEL 96
```

RESULT 13
AAW48211
ID AAW48211 standard; protein; 107 AA.
XX
AC AAW48211;
XX
DT 30-JUN-1998 (first entry)
XX
DE Conus radiatus conantokin.
XX
KW Conantokin; predatory cone snail; treatment; neurologic disorder;
KW psychiatric disorder; anticonvulsant; neuroprotective; analgesic;
KW HIV infection; ophthalmic indication; memory; learning defect;
KW cognitive defect.
XX
OS Conus radiatus.
XX
PN WO9803541-A1.
XX
PD 29-JAN-1998.
XX
PF 21-JUL-1997; 97WO-US012618.
XX
PR 22-JUL-1996; 96US-00684742.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Abogadie FC, Cruz LJ, Olivera BM, Walker C, Colledge C;
PI Hillyard DR, Jimenez E, Layer RT, Zhou L, Shen GS, Mccabe RT;
PI Rivier JE;
XX
DR WPI; 1998-120694/11.
DR N-PSDB; AAV20505.
XX
PT New conantokin peptide(s) - useful for e.g. treating neurologic or
PT psychiatric disorders, or the management of pain.
XX
PS Example 4; Page 81-82; 122pp; English.
XX
CC The present sequence is Conus radiatus conantokin, peptide derivatives of
CC which can be used to treat neurologic and psychiatric disorders, e.g. as
CC an anticonvulsant, neuroprotective or analgesic agent. Neurologic and
CC psychiatric disorders include epilepsy, convulsions, neurotoxic injury
CC (associated with conditions of hypoxia, anoxia or ischaemia, which
CC typically follow stroke, cerebrovascular accident, brain or spinal cord
CC trauma, myocardial infarct, physical trauma, drowning, suffocation,
CC perinatal asphyxia or hypoglycaemic events), neurodegeneration

CC (associated with Alzheimer's disease, senile dementia, Amyotrophic
CC Lateral Sclerosis, Multiple Sclerosis, Parkinson's disease, Huntington's
CC disease, Down's Syndrome, Korsakoff's disease, schizophrenia, AIDS
CC dementia, multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures), chemical toxicity (such as
CC addiction, and morphine, opiate, opioid and barbiturate tolerance), pain
CC (acute, chronic, migraine), anxiety, major depression, manic-depressive
CC illness, obsessive-compulsive disorder, schizophrenia and mood disorders
CC (such as bipolar disorder, unipolar depression, dysthymia and seasonal
CC effective disorder) and dystonia (movement disorder), sleep disorder,
CC muscle relaxation and urinary incontinence. The peptide can also be used
CC to treat HIV infection, ophthalmic indication and memory, learning or
CC cognitive defects
XX
SQ Sequence 107 AA;

Query Match 75.7%; Score 362.5; DB 2; Length 107;
Best Local Similarity 80.0%; Pred. No. 7e-39;
Matches 76; Conservative 3; Mismatches 13; Indels 3; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||| |||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||   ||    : :||   |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95
```

RESULT 14
AAW49990
ID AAW49990 standard; protein; 107 AA.
XX
AC AAW49990;
XX
DT 30-JUN-1998 (first entry)
XX
DE Conus radiatus conantokin.
XX
KW Conantokin; predatory cone snail; treatment; neurolcgic disorder;
KW psychiatric disorder; anticonvulsant; neuroprotective; analgesic;
KW HIV infection; ophthalmic indication; memory; learning defect;
KW cognitive defect.
XX
OS Conus radiatus.
XX
PN WO9803189-A1.
XX
PD 29-JAN-1998.
XX
PF 21-JUL-1997; 97WO-US012652.
XX
PR 22-JUL-1996; 96US-00684750.
PR 06-DEC-1996; 96US-00762377.
XX
PA (COGN-) COGNETIX INC.
XX
PI Mccabe RT, Zhou L, Layer RT;
XX
DR WPI; 1998-120469/11.
DR N-PSDB; AAV17132.
XX
PT Use of conantokin peptide(s) - for treating disorders involving excessive
PT excitation of nerve cells by excitatory amino acids or agonists of the N-
PT methyl-D-aspartate receptor.
XX
PS Example 4; Page 81-82; 122pp; English.
XX
CC The present sequence is Conus radiatus conantokin, peptide derivatives of
CC which can be used to treat neurologic and psychiatric disorders, e.g. as
CC an anticonvulsant, neuroprotective or analgesic agent. Neurologic and
CC psychiatric disorders include epilepsy, convulsions, neurotoxic injury
CC (associated with conditions of hypoxia, anoxia or ischaemia, which
CC typically follow stroke, cerebrovascular accident, brain or spinal cord
CC trauma, myocardial infarct, physical trauma, drowning, suffocation,
CC perinatal asphyxia or hypoglycaemic events), neurodegeneration
CC (associated with Alzheimer's disease, senile dementia, Amyotrophic
CC Lateral Sclerosis, Multiple Sclerosis, Parkinson's disease, Huntington's
CC disease, Down's Syndrome, Korsakoff's disease, schizophrenia, AIDS
CC dementia, multi-infarct dementia, Binswanger dementia and neuronal damage
CC associated with uncontrolled seizures), chemical toxicity (such as
CC addiction, and morphine, opiate, opioid and barbiturate tolerance), pain
CC (acute, chronic, migraine), anxiety, major depression, manic-depressive
CC illness, obsessive-compulsive disorder, schizophrenia and mood disorders
CC (such as bipolar disorder, unipolar depression, dysthymia and seasonal
CC effective disorder) and dystonia (movement disorder), sleep disorder,
CC muscle relaxation and urinary incontinence. The peptide can also be used
CC to treat HIV infection, ophthalmic indication and memory, learning or
CC cognitive defects
XX
SQ Sequence 107 AA;

Query Match 75.7%; Score 362.5; DB 2; Length 107;
Best Local Similarity 80.0%; Pred. No. 7e-39;
Matches 76; Conservative 3; Mismatches 13; Indels 3; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||| |||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||   ||    : :||   |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95
```

RESULT 15
AAG79045
ID AAG79045 standard; protein; 107 AA.
XX
AC AAG79045;
XX
DT 10-DEC-2001 (first entry)
XX
DE Amino acid sequence of conantokin R precursor protein.
XX
KW Conantokin; cone snail; nerve cell excitation; NMDA receptor; epilepsy;
KW N-methyl-D-aspartate receptor; pain; psychiatric disorder;
KW neutrotoxic injury; hypoxia; anoxia; ischemia; neurodegeneration;
KW chemical toxicity; addiction; drug craving; psychiatric disorder;
KW anxiety; depression; obsessive compulsive disorder; schizophrenia;
KW mood disorder; ophthalmic disorder; neurological disorder; dystonia;
KW sleep disorder; muscle relaxation; urinary incontinence;
KW cognition enhancement; HIV infection.
XX
OS Conus radiatus.
XX
PN US6277825-B1.
XX
PD 21-AUG-2001.
XX
PF 20-JUL-1999; 99US-00357141.
XX
PR 22-JUL-1996; 96US-00684750.
PR 06-DEC-1996; 96US-00762377.
PR 21-JUL-1997; 97WO-US012652.
PR 10-FEB-1999; 99US-00142076.
PR 01-APR-1999; 99US-00283277.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Olivera BM, Mcintosh JM, Mccabe RT, Layer RT, Zhou L;
XX
DR WPI; 2001-601377/68.
DR N-PSDB; AAI65369.
XX

PT Use of conantokin peptide or its derivatives or a conantokin peptide
PT chimera for treating disorders e.g. migraine.
XX
PS Example 4; Col 63-66; 60pp; English.
XX
CC The present sequence represents conantokin precursor protein. Conantokins
CC differ from conotoxins, in that they contain gamma-carboxyglutamic acid.
CC The conantokins are derived from the venom of cone snails. They are used
CC for the treatment of disorders in which the pathophysiology involves
CC excessive excitation of nerve cells by excitatory amino acids or agonist
CC of N-methyl-D-aspartate (NMDA) receptor. The conantokin peptides are used
CC for the treatment of disorders such as pain; neurologic or psychiatric
CC disorders such epilepsy; for reducing neutrotoxic injury associated with
CC conditions of hypoxia, anoxia or ischemia; for treating neurodegeneration
CC ; for treating chemical toxicity such as addiction, drug craving, alcohol
CC abuse, morphine, opioid and barbiturate tolerance; for treating
CC psychiatric disorders such as anxiety, major depression, manic-depression
CC illness, obsessive compulsive disorder, schizophrenia or mood disorder;
CC for treating ophthalmic disorder; for treating additional neurological
CC disorders e.g. dystonia, sleep disorder, muscle relaxation and urinary
CC incontinence; for memory/cognition enhancement; for treating HIV
CC infection
XX
SQ Sequence 107 AA;

Query Match 75.7%; Score 362.5; DB 4; Length 107;
Best Local Similarity 80.0%; Pred. No. 7e-39;
Matches 76; Conservative 3; Mismatches 13; Indels 3; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||| |||||||||||||||||||||||  ||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||   ||     : :||   |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95
```

Search completed: June 2, 2004, 18:09:44
Job time : 107.783 secs

GenCore version 5.1.6

DM protein - protein search, using sw model

Run on: June 2, 2004, 18:12:24 ; Search time 30.9302 Seconds
(without alignments)
158.565 Million cell updates/sec

Title: US-10-092-367-73
Perfect score: 479
Sequence: 1 MQLYTYLYLLVPLVTFYLIL.....GNMRGGEVRESAETLHEITP 95

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	362.5	75.7	107	3	US-09-142-078-50	Sequence 50, Appl
2	362.5	75.7	107	3	US-09-357-141-50	Sequence 50, Appl
3	362.5	75.7	107	4	US-09-533-889-50	Sequence 50, Appl
4	362.5	75.7	107	4	US-09-142-080-50	Sequence 50, Appl
5	353	73.7	100	3	US-09-142-078-46	Sequence 46, Appl
6	353	73.7	100	3	US-09-247-527-2	Sequence 2, Appli
7	353	73.7	100	3	US-09-357-141-46	Sequence 46, Appl
8	353	73.7	100	4	US-09-533-889-46	Sequence 46, Appl
9	353	73.7	100	4	US-09-142-080-46	Sequence 46, Appl
10	347	72.4	103	3	US-09-142-078-56	Sequence 56, Appl
11	347	72.4	103	3	US-09-357-141-56	Sequence 56, Appl
12	347	72.4	103	4	US-09-533-889-56	Sequence 56, Appl
13	347	72.4	103	4	US-09-142-080-56	Sequence 56, Appl
14	332	69.3	93	3	US-09-142-078-64	Sequence 64, Appl
15	332	69.3	93	3	US-09-357-141-64	Sequence 64, Appl
16	332	69.3	93	4	US-09-533-889-64	Sequence 64, Appl
17	332	69.3	93	4	US-09-142-080-64	Sequence 64, Appl
18	321	67.0	95	3	US-09-142-078-66	Sequence 66, Appl
19	321	67.0	95	3	US-09-357-141-66	Sequence 66, Appl
20	321	67.0	95	4	US-09-533-889-66	Sequence 66, Appl
21	321	67.0	95	4	US-09-142-080-66	Sequence 66, Appl
22	315.5	65.9	94	3	US-09-142-078-58	Sequence 58, Appl
23	315.5	65.9	94	3	US-09-357-141-58	Sequence 58, Appl
24	315.5	65.9	94	4	US-09-533-889-58	Sequence 58, Appl
25	315.5	65.9	94	4	US-09-142-080-58	Sequence 58, Appl
26	309	64.5	98	3	US-09-142-078-54	Sequence 54, Appl
27	309	64.5	98	3	US-09-357-141-54	Sequence 54, Appl
28	309	64.5	98	4	US-09-533-889-54	Sequence 54, Appl
29	309	64.5	98	4	US-09-142-080-54	Sequence 54, Appl
30	305	63.7	95	3	US-09-142-078-62	Sequence 62, Appl
31	305	63.7	95	3	US-09-357-141-62	Sequence 62, Appl
32	305	63.7	95	4	US-09-533-889-62	Sequence 62, Appl
33	305	63.7	95	4	US-09-142-080-62	Sequence 62, Appl
34	123	25.7	30	3	US-09-247-527-16	Sequence 16, Appl
35	90	18.8	20	3	US-09-247-527-17	Sequence 17, Appl
36	88	18.4	37	3	US-09-247-527-13	Sequence 13, Appl
37	84	17.5	26	3	US-09-247-527-19	Sequence 19, Appl
38	81	16.9	20	3	US-09-247-527-18	Sequence 18, Appl
39	79	16.5	328	4	US-09-065-040-12	Sequence 12, Appl
40	74	15.4	20	3	US-09-247-527-5	Sequence 5, Appli
41	74	15.4	20	3	US-09-247-527-15	Sequence 15, Appl
42	74	15.4	20	3	US-09-247-527-20	Sequence 20, Appl
43	74	15.4	25	3	US-09-247-527-14	Sequence 14, Appl
44	73	15.2	20	3	US-09-247-527-22	Sequence 22, Appl
45	73	15.2	20	3	US-09-247-527-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-142-078-50
; Sequence 50, Application US/09142078
; Patent No. 6172041
; GENERAL INFORMATION:
; APPLICANT: McCabe, R. Tyler
; APPLICANT: Zhou, Li-Ming
; APPLICANT: Layer, Richard T.
; TITLE OF INVENTION: Use of Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, p.c.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,078
; FILING DATE: 10-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12652
; FILING DATE: 21-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/762,377
; FILING DATE: 06-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/684,750
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-135.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-142-078-50

Query Match 75.7%; Score 362.5; DB 3; Length 107;

```
  Best Local Similarity   80.0%;  Pred. No. 6.6e-40;
  Matches   76;  Conservative    3;  Mismatches   13;  Indels    3;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||| ||||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||   ||    : :||   |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95


RESULT 2
US-09-357-141-50
; Sequence 50, Application US/09357141
; Patent No. 6277825
; GENERAL INFORMATION:
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  McCabe, R. Tyler
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Zhou, Li-Ming
;  TITLE OF INVENTION:  Use of Conantokins for Treating Pain
;  FILE REFERENCE:  2314-171
;  CURRENT APPLICATION NUMBER: US/09/357,141
;  CURRENT FILING DATE:  1999-07-20
;  PRIOR APPLICATION NUMBER:  US 09/283,277
;  PRIOR FILING DATE:  1999-04-01
;  PRIOR APPLICATION NUMBER:  US 09/142,078
;  PRIOR FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  WO US97/12652
;  PRIOR FILING DATE:  1997-07-21
;  PRIOR APPLICATION NUMBER:  US 08/762,377
;  PRIOR FILING DATE:  1996-12-06
;  PRIOR APPLICATION NUMBER:  US 08/684,750
;  PRIOR FILING DATE:  1996-07-22
;  NUMBER OF SEQ ID NOS:  71
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 50
;   LENGTH: 107
;   TYPE: PRT
;   ORGANISM: Conus radiatus
US-09-357-141-50

  Query Match             75.7%;  Score 362.5;  DB 3;  Length 107;
  Best Local Similarity   80.0%;  Pred. No. 6.6e-40;
  Matches   76;  Conservative    3;  Mismatches   13;  Indels    3;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||| ||||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||   ||    : :||   |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95


RESULT 3
US-09-533-889-50
; Sequence 50, Application US/09533889
; Patent No. 6399574
;  GENERAL INFORMATION:
;    APPLICANT:  McCabe, R. Tyler
;    APPLICANT:  Zhou, Li-Ming
;    APPLICANT:  Layer, Richard T.
;    APPLICANT:  Olivera, Baldomero M.
;    APPLICANT:  McIntosh, J. Michael
;    TITLE OF INVENTION:  Use of Conantokins
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, p.c.
;      STREET:  555 Thirteenth Street, N.W., Suite 701-E
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/533,889
;      FILING DATE:  22 MAR-2000
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/142,078
;      FILING DATE:  10-FEB-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US97/12652
;      FILING DATE:  21-JUL-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/762,377
;      FILING DATE:  06-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/684,750
;      FILING DATE:  22-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  2314-168.A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-783-6040
;      TELEFAX:  202-783-6031
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  107 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-533-889-50

  Query Match             75.7%;  Score 362.5;  DB 4;  Length 107;
  Best Local Similarity   80.0%;  Pred. No. 6.6e-40;
  Matches   76;  Conservative    3;  Mismatches   13;  Indels    3;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||| ||||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||   ||    : :||   |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95


RESULT 4
US-09-142-080-50
; Sequence 50, Application US/09142080
; Patent No. 6515103
;  GENERAL INFORMATION:
;       APPLICANT: Abogadie, Fe C.
;                  Cruz, Lourdes J.
;                  Olivera, Baldomero M.
;                  Walker, Craig
;                  Colledge, Clark
;                  Hillyard, David R.
;                  Jimenez, Elsie
;                  Layer, Richard T.
;                  Zhou, Li-Ming
;                  McCabe, R. Tyler
;       TITLE OF INVENTION: Conantokins
;       NUMBER OF SEQUENCES: 71
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, p.c.
;            STREET: 555 Thirteenth Street, N.W., Suite 701-E
;            CITY: Washington
;            STATE: D.C.
```

; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,080
; FILING DATE: 11-May-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12618
; FILING DATE: 21-JUL-1997
; APPLICATION NUMBER: US 08/684,742
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-134.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-142-080-50

Query Match 75.7%; Score 362.5; DB 4; Length 107;
Best Local Similarity 80.0%; Pred. No. 6.6e-40;
Matches 76; Conservative 3; Mismatches 13; Indels 3; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||| |||||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              | |||||| ||||||||    ||     : :||    |
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95
```

RESULT 5
US-09-142-078-46
; Sequence 46, Application US/09142078
; Patent No. 6172041
; GENERAL INFORMATION:
; APPLICANT: McCabe, R. Tyler
; APPLICANT: Zhou, Li-Ming
; APPLICANT: Layer, Richard T.
; TITLE OF INVENTION: Use of Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, p.c.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,078
; FILING DATE: 10-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12652
; FILING DATE: 21-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/762,377
; FILING DATE: 06-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/684,750
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-135.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-142-078-46

Query Match 73.7%; Score 353; DB 3; Length 100;
Best Local Similarity 79.8%; Pred. No. 1.1e-38;
Matches 75; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              | |||||||||||||||:||||||||  ||||||||  |||||||  ||||||||||||||||
Db          1 MHLYTYLYLLVPLVTFHLILGTGTLDDGGALTERRSADATALKAEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHE 92
              |||||||| ||||:|||    |  ||:|: | : |
Db         61 GKDRLTQMKRILKQRGNKARGEEEVQENQELIRE 94
```

RESULT 6
US-09-247-527-2
; Sequence 2, Application US/09247527A
; Patent No. 6197535
; GENERAL INFORMATION:
; APPLICANT: Bandyopadhyay, Pradip K.
; APPLICANT: Walker, Craig S.
; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: Methods for Purifying and Assaying a Conus
; TITLE OF INVENTION: Gamma-Carboxylase
; FILE REFERENCE: Conus Gamma-Carboxylase
; CURRENT APPLICATION NUMBER: US/09/247,527A
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 60/074,204
; EARLIER FILING DATE: 1998-02-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Conus geographus
US-09-247-527-2

Query Match 73.7%; Score 353; DB 3; Length 100;
Best Local Similarity 79.8%; Pred. No. 1.1e-38;
Matches 75; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              | |||||||||||||||:||||||||  ||||||||  |||||||  ||||||||||||||||
Db          1 MHLYTYLYLLVPLVTFHLILGTGTLDDGGALTERRSADATALKAEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHE 92
              |||||||| ||||:|||    |  ||:|: | : |
Db         61 GKDRLTQMKRILKQRGNKARGEEEVQENQELIRE 94
```

RESULT 7
US-09-357-141-46
; Sequence 46, Application US/09357141
; Patent No. 6277825

```
; GENERAL INFORMATION:
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  McCabe, R. Tyler
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Zhou, Li-Ming
;  TITLE OF INVENTION:  Use of Conantokins for Treating Pain
;  FILE REFERENCE:  2314-171
;  CURRENT APPLICATION NUMBER: US/09/357,141
;  CURRENT FILING DATE:  1999-07-20
;  PRIOR APPLICATION NUMBER:  US 09/283,277
;  PRIOR FILING DATE:  1999-04-01
;  PRIOR APPLICATION NUMBER:  US 09/142,078
;  PRIOR FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  WO US97/12652
;  PRIOR FILING DATE:  1997-07-21
;  PRIOR APPLICATION NUMBER:  US 08/762,377
;  PRIOR FILING DATE:  1996-12-06
;  PRIOR APPLICATION NUMBER:  US 08/684,750
;  PRIOR FILING DATE:  1996-07-22
;  NUMBER OF SEQ ID NOS:  71
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 46
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Conus geographus
US-09-357-141-46

  Query Match             73.7%;  Score 353;  DB 3;  Length 100;
  Best Local Similarity   79.8%;  Pred. No. 1.1e-38;
  Matches   75;  Conservative    5;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              | |||||||||||||||:||||||||  |||||||| ||||||| ||||||||||||||||
Db          1 MHLYTYLYLLVPLVTFHLILGTGTLDDGGALTERRSADATALKAEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHE 92
              |||||||| ||||:|||    |  ||:|: | : |
Db         61 GKDRLTQMKRILKQRGNKARGEEEVQENQELIRE 94


RESULT 8
US-09-533-889-46
; Sequence 46, Application US/09533889
; Patent No. 6399574
;  GENERAL INFORMATION:
;    APPLICANT:  McCabe, R. Tyler
;    APPLICANT:  Zhou, Li-Ming
,    APPLICANT:  Layer, Richard T.
,    APPLICANT:  Olivera, Baldomero M.
,    APPLICANT:  McIntosh, J. Michael
;    TITLE OF INVENTION:  Use of Conantokins
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
,      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, p.c.
,      STREET:  555 Thirteenth Street, N.W., Suite 701-E
,      CITY:  Washington
,      STATE:  D.C.
,      COUNTRY:  USA
,      ZIP:  20004
,    COMPUTER READABLE FORM:
,      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
,      OPERATING SYSTEM:  PC-DOS/MS-DOS
,      SOFTWARE:  PatentIn Release #1.0, Version #1.30
,    CURRENT APPLICATION DATA:
,      APPLICATION NUMBER:  US/09/533,889
,      FILING DATE:  22 MAR-2000
,    PRIOR APPLICATION DATA:
,      APPLICATION NUMBER:  US 09/142,078
,      FILING DATE:  10-FEB-1999
,    PRIOR APPLICATION DATA:
,      APPLICATION NUMBER:  WO US97/12652
;      FILING DATE:  21-JUL-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/762,377
;      FILING DATE:  06-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/684,750
;      FILING DATE:  22-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  2314-168.A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-783-6040
;      TELEFAX:  202-783-6031
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  100 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-533-889-46

  Query Match             73.7%;  Score 353;  DB 4;  Length 100;
  Best Local Similarity   79.8%;  Pred. No. 1.1e-38;
  Matches   75;  Conservative    5;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              | |||||||||||||||:||||||||  |||||||| ||||||| ||||||||||||||||
Db          1 MHLYTYLYLLVPLVTFHLILGTGTLDDGGALTERRSADATALKAEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHE 92
              |||||||| ||||:|||    |  ||:|: | : |
Db         61 GKDRLTQMKRILKQRGNKARGEEEVQENQELIRE 94


RESULT 9
US-09-142-080-46
; Sequence 46, Application US/09142080
; Patent No. 6515103
;   GENERAL INFORMATION:
;        APPLICANT: Abogadie, Fe C.
;                   Cruz, Lourdes J.
;                   Olivera, Baldomero M.
;                   Walker, Craig
;                   Colledge, Clark
;                   Hillyard, David R.
;                   Jimenez, Elsie
;                   Layer, Richard T.
;                   Zhou, Li-Ming
;                   McCabe, R. Tyler
;        TITLE OF INVENTION: Conantokins
;        NUMBER OF SEQUENCES: 71
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, p.c.
;             STREET: 555 Thirteenth Street, N.W., Suite 701-E
;             CITY: Washington
;             STATE: D.C.
;             COUNTRY: USA
;             ZIP: 20004
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/142,080
;             FILING DATE: 11-May-2000
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: WO US97/12618
;             FILING DATE: 21-JUL-1997
;             APPLICATION NUMBER: US 08/684,742
;             FILING DATE: 22-JUL-1996
;        ATTORNEY/AGENT INFORMATION:
```

; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-134.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-142-080-46

Query Match 73.7%; Score 353; DB 4; Length 100;
Best Local Similarity 79.8%; Pred. No. 1.1e-38;
Matches 75; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              | |||||||||||||||:||||||| |||||||| |||||||||||||||||||||||||
Db          1 MHLYTYLYLLVPLVTFHLILGTGTLDDGGALTERRSADATALKAEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHE 92
              |||||||| ||||:|||   |  ||:|: | : |
Db         61 GKDRLTQMKRILKQRGNKARGEEEVQENQELIRE 94
```

RESULT 10
US-09-142-078-56
; Sequence 56, Application US/09142078
; Patent No. 6172041
; GENERAL INFORMATION:
; APPLICANT: McCabe, R. Tyler
; APPLICANT: Zhou, Li-Ming
; APPLICANT: Layer, Richard T.
; TITLE OF INVENTION: Use of Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, p.c.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,078
; FILING DATE: 10-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12652
; FILING DATE: 21-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/762,377
; FILING DATE: 06-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/684,750
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-135.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-142-078-56

Query Match 72.4%; Score 347; DB 3; Length 103;
Best Local Similarity 89.2%; Pred. No. 6.8e-38;
Matches 74; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||:||||||| |||||||||  |||||||||| ||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSTDATALKPEPV-LQKSAARSTDDN 59

Qy         61 GKDRLTQMIRILKKRG-NMRGGE 82
              |||||||| ||||||| | || |
Db         60 GKDRLTQMKRILKKRGKNARGDE 82
```

RESULT 11
US-09-357-141-56
; Sequence 56, Application US/09357141
; Patent No. 6277825
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: McCabe, R. Tyler
; APPLICANT: Layer, Richard T.
; APPLICANT: Zhou, Li-Ming
; TITLE OF INVENTION: Use of Conantokins for Treating Pain
; FILE REFERENCE: 2314-171
; CURRENT APPLICATION NUMBER: US/09/357,141
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 09/283,277
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: US 09/142,078
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: WO US97/12652
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: US 08/762,377
; PRIOR FILING DATE: 1996-12-06
; PRIOR APPLICATION NUMBER: US 08/684,750
; PRIOR FILING DATE: 1996-07-22
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Conus sulcatus
US-09-357-141-56

Query Match 72.4%; Score 347; DB 3; Length 103;
Best Local Similarity 89.2%; Pred. No. 6.8e-38;
Matches 74; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||:||||||| |||||||||  |||||||||| ||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSTDATALKPEPV-LQKSAARSTDDN 59

Qy         61 GKDRLTQMIRILKKRG-NMRGGE 82
              |||||||| ||||||| | || |
Db         60 GKDRLTQMKRILKKRGKNARGDE 82
```

RESULT 12
US-09-533-889-56
; Sequence 56, Application US/09533889
; Patent No. 6399574
; GENERAL INFORMATION:
; APPLICANT: McCabe, R. Tyler
; APPLICANT: Zhou, Li-Ming
; APPLICANT: Layer, Richard T.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; TITLE OF INVENTION: Use of Conantokins

```
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, p.c.
;      STREET:  555 Thirteenth Street, N.W., Suite 701-E
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/533,889
;      FILING DATE:  22 MAR-2000
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/142,078
;      FILING DATE:  10-FEB-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US97/12652
;      FILING DATE:  21-JUL-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/762,377
;      FILING DATE:  06-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/684,750
;      FILING DATE:  22-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  2314-168.A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-783-6040
;      TELEFAX:  202-783-6031
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  103 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-533-889-56

  Query Match             72.4%;  Score 347;  DB 4;  Length 103;
  Best Local Similarity   89.2%;  Pred. No. 6.8e-38;
  Matches   74;  Conservative    1;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||:|||||||| |||||||||  |||||||||| ||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSTDATALKPEPV-LQKSAARSTDDN 59

Qy         61 GKDRLTQMIRILKKRG-NMRGGE 82
              |||||||| ||||||| | || |
Db         60 GKDRLTQMKRILKKRGKNARGDE 82


RESULT 13
US-09-142-080-56
; Sequence 56, Application US/09142080
; Patent No. 6515103
;  GENERAL INFORMATION:
;       APPLICANT: Abogadie, Fe C.
;                  Cruz, Lourdes J.
;                  Olivera, Baldomero M.
;                  Walker, Craig
;                  Colledge, Clark
;                  Hillyard, David R.
;                  Jimenez, Elsie
;                  Layer, Richard T.
;                  Zhou, Li-Ming
;                  McCabe, R. Tyler
;       TITLE OF INVENTION: Conantokins
;       NUMBER OF SEQUENCES: 71
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, p.c.
;            STREET: 555 Thirteenth Street, N.W., Suite 701-E
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: USA
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/142,080
;            FILING DATE: 11-May-2000
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: WO US97/12618
;            FILING DATE: 21-JUL-1997
;            APPLICATION NUMBER: US 08/684,742
;            FILING DATE: 22-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ihnen, Jeffrey L.
;            REGISTRATION NUMBER: 28,957
;            REFERENCE/DOCKET NUMBER: 2314-134.A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 202-783-6040
;            TELEFAX: 202-783-6031
;   INFORMATION FOR SEQ ID NO: 56:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 103 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-09-142-080-56

  Query Match             72.4%;  Score 347;  DB 4;  Length 103;
  Best Local Similarity   89.2%;  Pred. No. 6.8e-38;
  Matches   74;  Conservative    1;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||:|||||||| |||||||||  |||||||||| ||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSTDATALKPEPV-LQKSAARSTDDN 59

Qy         61 GKDRLTQMIRILKKRG-NMRGGE 82
              |||||||| ||||||| | || |
Db         60 GKDRLTQMKRILKKRGKNARGDE 82


RESULT 14
US-09-142-078-64
; Sequence 64, Application US/09142078
; Patent No. 6172041
;  GENERAL INFORMATION:
;    APPLICANT:  McCabe, R. Tyler
;    APPLICANT:  Zhou, Li-Ming
;    APPLICANT:  Layer, Richard T.
;    TITLE OF INVENTION:  Use of Conantokins
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, p.c.
;      STREET:  555 Thirteenth Street, N.W., Suite 701-E
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/142,078
```

; FILING DATE: 10-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12652
; FILING DATE: 21-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/762,377
; FILING DATE: 06-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/684,750
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-135.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-142-078-64

Query Match 69.3%; Score 332; DB 3; Length 93;
Best Local Similarity 80.7%; Pred. No. 5.5e-36;
Matches 71; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

```
Qy          8 YLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQ 67
              ||||||| |:|||||||| || |||||| ||||||||||||||||||||||||||||||
Db          1 YLLVPLVAFHLILGTGTLAHGDALTERRSADATALKPEPVLLQKSAARSTDDNGKDRLTQ 60

Qy         68 MIRILKKRGNMRGG--EVRESAETLHEI 93
                | ||||||| |  | || |||: |:
Db         61 RKRTLKKRGNMARGYEEDREIAETVREL 88
```

RESULT 15
US-09-357-141-64
; Sequence 64, Application US/09357141
; Patent No. 6277825
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
, APPLICANT: McCabe, R. Tyler
, APPLICANT: Layer, Richard T.
, APPLICANT: Zhou, Li-Ming
; TITLE OF INVENTION: Use of Conantokins for Treating Pain
; FILE REFERENCE: 2314-171
; CURRENT APPLICATION NUMBER: US/09/357,141
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 09/283,277
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: US 09/142,078
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: WO US97/12652
; PRIOR FILING DATE: 1997-07-21
, PRIOR APPLICATION NUMBER: US 08/762,377
, PRIOR FILING DATE: 1996-12-06
, PRIOR APPLICATION NUMBER: US 08/684,750
, PRIOR FILING DATE: 1996-07-22
; NUMBER OF SEQ ID NOS: 71
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 64
, LENGTH: 93
, TYPE: PRT
, ORGANISM: Conus caracteristicus
US-09-357-141-64

Query Match 69.3%; Score 332; DB 3; Length 93;
Best Local Similarity 80.7%; Pred. No. 5.5e-36;
Matches 71; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

```
Qy          8 YLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQ 67
              ||||||| |:|||||||| || |||||| ||||||||||||||||||||||||||||||
Db          1 YLLVPLVAFHLILGTGTLAHGDALTERRSADATALKPEPVLLQKSAARSTDDNGKDRLTQ 60

Qy         68 MIRILKKRGNMRGG--EVRESAETLHEI 93
                | ||||||| |  | || |||: |:
Db         61 RKRTLKKRGNMARGYEEDREIAETVREL 88
```

Search completed: June 2, 2004, 18:13:55
Job time : 30.9302 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:13:14 ; Search time 80.2713 Seconds
(without alignments)
332.960 Million cell updates/sec

Title: US-10-092-367-73
Perfect score: 479
Sequence: 1 MQLYTYLYLLVPLVTFYLIL.....GNMRGGEVRESAETLHEITP 95

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	479	100.0	95	12	US-10-092-367-73	Sequence 73, Appl
2	471	98.3	95	12	US-10-092-367-64	Sequence 64, Appl
3	469	97.9	95	12	US-10-092-367-61	Sequence 61, Appl
4	436	91.0	97	12	US-10-092-367-67	Sequence 67, Appl
5	405.5	84.7	94	12	US-10-092-367-103	Sequence 103, App
6	385	80.4	100	12	US-10-092-367-91	Sequence 91, Appl
7	379	79.1	98	12	US-10-092-367-70	Sequence 70, Appl
8	377	78.7	101	14	US-10-207-780-57	Sequence 57, Appl
9	371	77.5	101	12	US-10-092-367-106	Sequence 106, App
10	364.5	76.1	102	12	US-10-092-367-40	Sequence 40, Appl
11	363	75.8	107	12	US-10-092-367-52	Sequence 52, Appl
12	362.5	75.7	102	14	US-10-207-780-61	Sequence 61, Appl
13	362.5	75.7	107	14	US-10-357-467-50	Sequence 50, Appl
14	361.5	75.5	102	14	US-10-207-780-59	Sequence 59, Appl
15	359	74.9	96	12	US-10-092-367-79	Sequence 79, Appl
16	359	74.9	99	12	US-10-092-367-100	Sequence 100, App
17	358.5	74.8	102	12	US-10-092-367-37	Sequence 37, Appl
18	356	74.3	99	12	US-10-092-367-82	Sequence 82, Appl
19	354	73.9	96	12	US-10-092-367-127	Sequence 127, App
20	353	73.7	100	14	US-10-207-780-47	Sequence 47, Appl
21	353	73.7	100	14	US-10-357-467-46	Sequence 46, Appl
22	352.5	73.6	102	12	US-10-092-367-43	Sequence 43, Appl
23	352	73.5	99	12	US-10-092-367-88	Sequence 88, Appl
24	350.5	73.2	101	12	US-10-092-367-55	Sequence 55, Appl
25	350	73.1	102	12	US-10-092-367-115	Sequence 115, App
26	349	72.9	98	14	US-10-207-780-73	Sequence 73, Appl
27	348	72.7	102	12	US-10-092-367-121	Sequence 121, App
28	347.5	72.5	103	12	US-10-092-367-58	Sequence 58, Appl
29	347.5	72.5	114	12	US-10-092-367-94	Sequence 94, Appl
30	347.5	72.5	114	12	US-10-092-367-97	Sequence 97, Appl
31	347	72.4	103	14	US-10-357-467-56	Sequence 56, Appl
32	345	72.0	96	14	US-10-207-780-75	Sequence 75, Appl
33	343	71.6	101	12	US-10-092-367-85	Sequence 85, Appl
34	343	71.6	111	14	US-10-207-780-87	Sequence 87, Appl
35	343	71.6	111	14	US-10-207-780-89	Sequence 89, Appl
36	342	71.4	98	14	US-10-207-780-77	Sequence 77, Appl
37	341	71.2	100	14	US-10-207-780-55	Sequence 55, Appl
38	340.5	71.1	97	12	US-10-092-367-130	Sequence 130, App
39	336	70.1	104	14	US-10-207-780-91	Sequence 91, Appl
40	335.5	70.0	101	12	US-10-092-367-118	Sequence 118, App
41	335	69.9	100	12	US-10-092-367-46	Sequence 46, Appl
42	335	69.9	100	14	US-10-207-780-53	Sequence 53, Appl
43	335	69.9	102	12	US-10-092-367-124	Sequence 124, App
44	333	69.5	100	12	US-10-092-367-49	Sequence 49, Appl
45	332	69.3	93	14	US-10-357-467-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-10-092-367-73
; Sequence 73, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-73

Query Match 100.0%; Score 479; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.4e-50;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
              |||||||||||||||||||||||||||||||||||
Db         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
```

RESULT 2
US-10-092-367-64
; Sequence 64, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 64
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-64

Query Match 98.3%; Score 471; DB 12; Length 95;
Best Local Similarity 98.9%; Pred. No. 1.4e-49;
Matches 94; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
              |||||||||||||||||| ||||||||||||||
Db         61 GKDRLTQMIRILKKRGNMRGEEVRESAETLHEITP 95
```

RESULT 3
US-10-092-367-61
; Sequence 61, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 61
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-61

Query Match 97.9%; Score 469; DB 12; Length 95;
Best Local Similarity 97.9%; Pred. No. 2.4e-49;
Matches 93; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
              |||||||||||||||||| |||||||||||:||
Db         61 GKDRLTQMIRILKKRGNMRGEEVRESAETLHELTP 95
```

RESULT 4
US-10-092-367-67
; Sequence 67, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-67

Query Match 91.0%; Score 436; DB 12; Length 97;
Best Local Similarity 90.7%; Pred. No. 2.6e-45;
Matches 88; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||:|:||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPKPILLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNM--RGGEVRESAETLHEITP 95
              |||||||||||||||||    | ||||:||||:|:||
Db         61 GKDRLTQMIRILKKRGNMGRDGEEVREAAETLNELTP 97
```

RESULT 5
US-10-092-367-103
; Sequence 103, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103
; LENGTH: 94
; TYPE: PRT

; ORGANISM: Conus figulinus
US-10-092-367-103

Query Match 84.7%; Score 405.5; DB 12; Length 94;
Best Local Similarity 88.3%; Pred. No. 1.4e-41;
Matches 83; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEIT 94
              ||||||||   :|||||    ||||:||||||::
Db         61 GKDRLTQMKGTVKKRGN-TAEEVREAAETLHELS 93
```

RESULT 6
US-10-092-367-91
; Sequence 91, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Conus figulinus
US-10-092-367-91

Query Match 80.4%; Score 385; DB 12; Length 100;
Best Local Similarity 90.8%; Pred. No. 4.7e-39;
Matches 79; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESA 87
              ||||||| || ||||| |  || | |
Db         61 DKDRLTQMKRIFKKRGNKREEEVAEMA 87
```

RESULT 7
US-10-092-367-70
; Sequence 70, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 70
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Conus betulinus
US-10-092-367-70

Query Match 79.1%; Score 379; DB 12; Length 98;
Best Local Similarity 83.9%; Pred. No. 2.5e-38;
Matches 78; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||:|||||||||||||||| | |||||||| ||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLGHGGALTESRSADATALKPGPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEI 93
              |||||||| | |||||| |  : || |||: |:
Db         61 GKDRLTQMKRTLKKRGNTRYEDDREIAETVREL 93
```

RESULT 8
US-10-207-780-57
; Sequence 57, Application US/10207780
; Publication No. US20030144210A1
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; TITLE OF INVENTION: Gamma-Carboxyglutamate Containing Conopeptides
; FILE REFERENCE: Gla-Conopeptides
; CURRENT APPLICATION NUMBER: US/10/207,780
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: US/09/658,603
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/219,673
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/153,034
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Conus obscurus
US-10-207-780-57

Query Match 78.7%; Score 377; DB 14; Length 101;
Best Local Similarity 83.2%; Pred. No. 4.5e-38;
Matches 79; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              ||||||||||||||||:|||||||| |||||||| ||||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGN--MRGGEVRESAETLHEI 93
              ||||||||  |||||||    |  |:||  ||: |:
Db         61 GKDRLTQMKGILKKRGNTARRDEELREDVETILEL 95
```

RESULT 9
US-10-092-367-106
; Sequence 106, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation

; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Conus figulinus
US-10-092-367-106

Query Match 77.5%; Score 371; DB 12; Length 101;
Best Local Similarity 82.1%; Pred. No. 2.4e-37;
Matches 78; Conservative 6; Mismatches 9; Indels 2; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||| |
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDVN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHEI 93
              ||||||:| |||||||::  |    || || : |:
Db         61 GKDRLTEMKRILKKRGSISMGFEHRRREIAELVREL 95
```

RESULT 10
US-10-092-367-40
; Sequence 40, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Conus catus
US-10-092-367-40

Query Match 76.1%; Score 364.5; DB 12; Length 102;
Best Local Similarity 78.1%; Pred. No. 1.5e-36;
Matches 75; Conservative 8; Mismatches 10; Indels 3; Gaps 1;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:|||||||| |||||||||  |||||:|||||||||||||||:
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSGDATALRPEPVLLQKSAARSTDDS 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHEI 93
              |||||||| |||||:||   |:   :||  ||: |:
Db         61 GKDRLTQMKRILKKQGNTAKGDEELLREDVETVLEL 96
```

RESULT 11
US-10-092-367-52
; Sequence 52, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
; FILE REFERENCE: 2314-224-II
; CURRENT APPLICATION NUMBER: US/10/092,367
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/273,639
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Conus catus
US-10-092-367-52

Query Match 75.8%; Score 363; DB 12; Length 107;
Best Local Similarity 81.5%; Pred. No. 2.5e-36;
Matches 75; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:|||||||| ||||||||| |||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHE 92
              |||||| | |||||| | |  ||    | : :
Db         61 GKDRLTHMKRILKKRANKREPEVGSIPEAVRQ 92
```

RESULT 12
US-10-207-780-61
; Sequence 61, Application US/10207780
; Publication No. US20030144210A1
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Walker, Craig S.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; TITLE OF INVENTION: Gamma-Carboxyglutamate Containing Conopeptides
; FILE REFERENCE: Gla-Conopeptides
; CURRENT APPLICATION NUMBER: US/10/207,780
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: US/09/658,603
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/219,673
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/153,034
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Conus obscurus
US-10-207-780-61

Query Match 75.7%; Score 362.5; DB 14; Length 102;

```
  Best Local Similarity   81.2%;  Pred. No. 2.7e-36;
  Matches   78;  Conservative    5;  Mismatches   10;  Indels    3;  Gaps    2;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:||||||| |||||||| |||||||||||||||||||||||||
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGN--MRGGE-VRESAETLHEI 93
              |||||||  ||||:||   |  | :||  ||: |:
Db         61 GKDRLTQMKGILKKQGNTARRDEELLREDVETILEL 96


RESULT 13
US-10-357-467-50
; Sequence 50, Application US/10357467
; Publication No. US20030194729A1
;   GENERAL INFORMATION:
;        APPLICANT: Abogadie, Fe C.
;                   Cruz, Lourdes J.
;                   Olivera, Baldomero M.
;                   Walker, Craig
;                   Colledge, Clark
;                   Hillyard, David R.
;                   Jimenez, Elsie
;        TITLE OF INVENTION: Conantokins
;        NUMBER OF SEQUENCES: 71
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, p.c.
;             STREET: 1425 K Street, N.W., Suite 800
;             CITY: Washington
;             STATE: D.C.
;             COUNTRY: USA
;             ZIP: 20005
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/357,467
;             FILING DATE: 04-Feb-2003
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 09/142,080
;             FILING DATE: 15-MAY-2000
;             APPLICATION NUMBER: WO US97/12618
;             FILING DATE: 21-JUL-1997
;             APPLICATION NUMBER: US 08/684,742
;             FILING DATE: 22-JUL-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ihnen, Jeffrey L.
;             REGISTRATION NUMBER: 28,957
;             REFERENCE/DOCKET NUMBER: 2314-256.A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-783-6040
;             TELEFAX: 202-783-6031
;   INFORMATION FOR SEQ ID NO: 50:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 107 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-357-467-50

  Query Match             75.7%;  Score 362.5;  DB 14;  Length 107;
  Best Local Similarity   80.0%;  Pred. No. 2.9e-36;
  Matches   76;  Conservative    3;  Mismatches   13;  Indels    3;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||| |||||||||||||||||||||  |||||||||||||||:|||||||
Db          1 MQLYTYLYLLVSLVTFYLILGTGTLGHGGALTERRSTDATALKPEPVLLQKSSARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHE 92
              |||||||| |||||||||||||   |||||||||||
Db         61 GNDRLTQMKRILKKRGNKARGEEEVAKMAAELARE 95


RESULT 14
US-10-207-780-59
; Sequence 59, Application US/10207780
; Publication No. US20030144210A1
; GENERAL INFORMATION:
;  APPLICANT: Olivera, Baldomero M.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  APPLICANT:  University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  TITLE OF INVENTION: Gamma-Carboxyglutamate Containing Conopeptides
;  FILE REFERENCE: Gla-Conopeptides
;  CURRENT APPLICATION NUMBER: US/10/207,780
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: US/09/658,603
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: US 60/219,673
;  PRIOR FILING DATE: 2000-07-21
;  PRIOR APPLICATION NUMBER: US 60/153,034
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 99
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
;   LENGTH: 102
;   TYPE: PRT
;   ORGANISM: Conus obscurus
US-10-207-780-59

  Query Match             75.5%;  Score 361.5;  DB 14;  Length 102;
  Best Local Similarity   78.1%;  Pred. No. 3.6e-36;
  Matches   75;  Conservative    7;  Mismatches   11;  Indels    3;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||||||||||||:||||||| |||||||||  |||||||||||||||||||||:
Db          1 MQLYTYLYLLVPLVTFHLILGTGTLDHGGALTERRSGDATALKPEPVLLQKSAARSTDDS 60

Qy         61 GKDRLTQMIRILKKRGNMRGGE---VRESAETLHEI 93
              |||||||| |||||:||    :   :||  ||: |:
Db         61 GKDRLTQMKRILKKQGNTAKSDEELLREDVETVLEL 96


RESULT 15
US-10-092-367-79
; Sequence 79, Application US/10092367
; Publication No. US20030065138A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Walker, Craig S.
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Linear Gamma-Carboxyglutamate Rich Conotoxins
;  FILE REFERENCE: 2314-224-II
;  CURRENT APPLICATION NUMBER: US/10/092,367
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/273,639
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79
;   LENGTH: 96
;   TYPE: PRT
;   ORGANISM: Conus episcopatus
```

US-10-092-367-79

Query Match 74.9%; Score 359; DB 12; Length 96;
Best Local Similarity 79.6%; Pred. No. 6.7e-36;
Matches 74; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

```
Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |||||||  ||||||||||||||||| |||||||  | ||||||||||||||||||||||
Db          1 MQLYTYLCLLVPLVTFYLILGTGTLAHGGALTEHRSADATALKPEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGGEVRESAETLHEI 93
              ||||||:   ||||||| || :: |:   | :|
Db         61 GKDRLTRWKGILKKRGNTRGKDIVETITELEKI 93
```

Search completed: June 2, 2004, 18:15:58
Job time : 81.2713 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:10:29 ; Search time 25.7752 Seconds
(without alignments)
354.534 Million cell updates/sec

Title: US-10-092-367-73
Perfect score: 479
Sequence: 1 MQLYTYLYLLVPLVTFYLIL.....GNMRGGEVRESAETLHEITP 95

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72.5	15.1	650	2	A69390	3-hydroxyacyl-CoA
2	71.5	14.9	663	2	F90423	hypothetical prote
3	70.5	14.7	677	2	S33814	kinesin light chai
4	70.5	14.7	686	2	S33815	kinesin light chai
5	70	14.6	593	2	S26696	alkaline proteinas
6	70	14.6	755	2	S32103	filensin - bovine
7	70	14.6	1291	2	T06692	hypothetical prote
8	69.5	14.5	649	2	S33813	kinesin light chai
9	69.5	14.5	1298	2	T17199	CL3BB protein - ra
10	69.5	14.5	1341	2	T17200	CL3BC protein - ra
11	69.5	14.5	1527	2	T17198	CL3BA protein - ra
12	69.5	14.5	1550	2	T14327	alpha-latrotoxin r
13	69.5	14.5	1556	2	F96587	hypothetical prote
14	68.5	14.3	1583	2	T00727	myosin heavy chain
15	67.5	14.1	399	2	G69868	hypothetical prote
16	67.5	14.1	524	2	AF3096	conserved hypothet
17	67.5	14.1	524	2	C98190	hypothetical prote
18	67.5	14.1	1299	2	T18398	latrophilin-3, spl
19	67.5	14.1	1308	2	T18408	latrophilin-3, spl
20	67.5	14.1	1342	2	T18405	latrophilin-3, spl
21	67.5	14.1	1351	2	T18409	latrophilin-3, spl
22	67.5	14.1	1502	2	D84587	probable myosin he
23	67.5	14.1	1571	2	T18395	latrophilin-3, spl
24	67.5	14.1	1580	2	T18407	latrophilin-3, spl
25	67	14.0	230	2	F83273	probable hydrolase
26	66.5	13.9	245	1	C65210	hypothetical 26.3
27	66.5	13.9	245	2	C91255	hypothetical prote
28	66.5	13.9	245	2	G86095	hypothetical prote
29	66.5	13.9	248	2	G69352	branched-chain ami
30	66.5	13.9	513	2	AF3575	exodeoxyribonuclea
31	66.5	13.9	723	2	H82035	fatty oxidation co
32	66.5	13.9	1477	2	T00957	myosin heavy chain
33	66	13.8	929	2	I51027	type XII collagen
34	66	13.8	1579	2	T30516	type II DNA topois
35	65.5	13.7	578	2	H75256	malate oxidoreduct
36	65.5	13.7	2082	2	T37056	probable multi-dom
37	65	13.6	304	2	AE1069	mrr restriction sy
38	64.5	13.5	153	2	AB1907	hypothetical prote
39	64.5	13.5	231	2	AH1559	E. coli copper hom
40	64.5	13.5	328	2	G83363	asparaginase (EC 3
41	64.5	13.5	558	2	T09976	H+-transporting tw
42	64.5	13.5	922	2	T18878	hypothetical prote
43	64	13.4	151	2	F69212	autotrophic growth
44	64	13.4	157	1	RDECD7	dihydrofolate redu
45	64	13.4	157	2	S11706	dihydrofolate redu

ALIGNMENTS

RESULT 1
A69390
3-hydroxyacyl-CoA dehydrogenase (hbd-5) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 28-Jul-2000
C;Accession: A69390
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69390
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-650 <KLE>
A;Cross-references: GB:AE001026; GB:AE000782; NID:g2689349; PIDN:AAB90118.1; PID:g26494
C;Superfamily: probable 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenas
F;1-277/Domain: 3-hydroxyacyl-CoA dehydrogenase homology <HCD>
F;418-570/Domain: enoyl-CoA hydratase homology <ECH>

Query Match 15.1%; Score 72.5; DB 2; Length 650;
Best Local Similarity 22.8%; Pred. No. 10;
Matches 23; Conservative 16; Mismatches 29; Indels 33; Gaps 3;

```
Qy         18 LILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGN 77
              |::| ||:|||        :|:  ||     |:|     :   |   |:   :| |:::|
Db          4 LVVGAGTMGHG-------IAEVCALAGNEVIL-CDINENILKNALSRIEWSVRKLEQKGR 55

Qy         78 MRGGE-------------------------VRESAETLHEI 93
              ::| :                        | |  |  ||:
Db         56 IKGADDVLKRLKTTTDLVEAAKEADFVIEAVVEKTEVKHEV 96
```

RESULT 2
F90423
hypothetical protein SSO2514 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: F90423
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90423
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-663 <KUR>

A;Cross-references: GB:AE006641; NID:g13815814; PIDN:AAK42645.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO2514
C;Superfamily: probable 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase

Query Match 14.9%; Score 71.5; DB 2; Length 663;
Best Local Similarity 26.0%; Pred. No. 13;
Matches 19; Conservative 19; Mismatches 22; Indels 13; Gaps 3;

```
Qy         18 LILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGN 77
              |::| ||:|||        :|:  |:     | |     ::   :|  :|:   :  |::|
Db         10 LVVGAGTMGHG-------IAEIAAISGYKVYL-SDISQDILNNALERIRWSLSKLQER-- 59

Qy         78 MRGGEVRESAETL 90
                |:::|| :|:
Db         60 ---GQIKESVDTI 69
```

RESULT 3
S33814
kinesin light chain - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C;Accession: S33814
R;Wedaman, K.P.; Knight, A.E.; Kendrick-Jones, J.; Scholey, J.M.
J. Mol. Biol. 231, 155-158, 1993
A;Title: Sequences of sea urchin kinesin light chain isoforms.
A;Reference number: S33813; MUID:93267648; PMID:8496962
A;Accession: S33814
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-677 <WED>
A;Cross-references: EMBL:L10234; NID:g161527; PID:g161528

Query Match 14.7%; Score 70.5; DB 2; Length 677;
Best Local Similarity 27.8%; Pred. No. 17;
Matches 20; Conservative 16; Mismatches 31; Indels 5; Gaps 2;

```
Qy          8 YLLVPLVTFYLIL--GTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRL 65
              |: :|    ::::  | | |   |:|:: |   |:  :    || |   | :||:|  :
Db        578 YVEIPRSPPHVLVENGDGKLRRSGSLSKLR---ASVRRSSTKLLNKLKGRESDDDGGMKR 634

Qy         66 TQMIRILKKRGN 77
                : :|  |||
Db        635 ASSMSVLPSRGN 646
```

RESULT 4
S33815
kinesin light chain isoform 3 - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 08-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 01-Dec-2000
C;Accession: S33815; S36727
R;Wedaman, K.P.; Knight, A.E.; Kendrick-Jones, J.; Scholey, J.M.
J. Mol. Biol. 231, 155-158, 1993
A;Title: Sequences of sea urchin kinesin light chain isoforms.
A;Reference number: S33813; MUID:93267648; PMID:8496962
A;Accession: S33815
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-686 <WED>
A;Cross-references: EMBL:L10235
R;Wedaman, K.P.; Knight, A.E.; Kendrick-Jones, J.; Scholey, J.M.
submitted to the EMBL Data Library, February 1993
A;Reference number: S36727
A;Accession: S36727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-568,'A',570-686 <WE2>
A;Cross-references: EMBL:L10235; NID:g161529; PID:g161530

Query Match 14.7%; Score 70.5; DB 2; Length 686;
Best Local Similarity 27.8%; Pred. No. 18;
Matches 20; Conservative 16; Mismatches 31; Indels 5; Gaps 2;

```
Qy          8 YLLVPLVTFYLIL--GTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRL 65
              |: :|    ::::  | | |   |:|:: |   |:  :    || |   | :||:|  :
Db        587 YVEIPRSPPHVLVENGDGKLRRSGSLSKLR---ASVRRSSTKLLNKLKGRESDDDGGMKR 643

Qy         66 TQMIRILKKRGN 77
                : :|  |||
Db        644 ASSMSVLPSRGN 655
```

RESULT 5
S26696
alkaline proteinase secretion protein AprD PA1246 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
C;Accession: S26696; F83489
R;Duong, F.; Lazdunski, A.; Cami, B.; Murgier, M.
Gene 121, 47-54, 1992
A;Title: Sequence of a cluster of genes controlling synthesis and secretion of alkaline
A;Reference number: S26696; MUID:93051361; PMID:1427098
A;Accession: S26696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-593 <DUO>
A;Cross-references: EMBL:X64558; NID:g45279; PIDN:CAA45855.1; PID:g45280
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; E
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83489
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-593 <STO>
A;Cross-references: GB:AE004554; GB:AE004091; NID:g9947174; PIDN:AAG04635.1; GSPDB:GN00
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: aprD; PA1246
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C;Keywords: ATP; nucleotide binding; P-loop
F;349-543/Domain: ATP-binding cassette homology <ABC>
F;366-373/Region: nucleotide-binding motif A (P-loop)

Query Match 14.6%; Score 70; DB 2; Length 593;
Best Local Similarity 37.3%; Pred. No. 17;
Matches 22; Conservative 7; Mismatches 28; Indels 2; Gaps 2;

```
Qy         19 ILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARST-DDNGKDRLTQMIRILKKRG 76
              :|| |  |  |    :|:| | ||   | |:      |  ||:|:  |   |: || ||
Db        461 VLGVGGAGLSGG-QRQRIALARALYGAPTLVVLDEPNSNLDDSGEQALLAAIQALKARG 518
```

RESULT 6
S32103
filensin - bovine
N;Alternate names: intermediate filament protein
C;Species: Bos primigenius taurus (cattle)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C;Accession: S32103; A40690
R;Gounari, F.; Merdes, A.; Quinlan, R.; Hess, J.; FitzGerald, P.G.; Ouzounis, C.; Georga
submitted to the EMBL Data Library, March 1993
A;Description: Bovine filensin possesses primary and secondary structure similarity to i
A;Reference number: S32103
A;Accession: S32103
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-755 <GOU>
A;Cross-references: EMBL:X72388; NID:g287751; PID:g287752
R;Gounari, F.; Merdes, A.; Quinlan, R.; Hess, J.; FitzGerald, P.G.; Ouzounis, C.A.; Geor
J. Cell Biol. 121, 847-853, 1993
A;Title: Bovine filensin possesses primary and secondary structure similarity to interme

```
A;Reference number: A40690; MUID:93260017; PMID:8491777
A;Accession: A40690
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-622,'RP',625-755 <GO2>
A;Experimental source: lens
A;Note: sequence extracted from NCBI backbone (NCBIN:132495, NCBIP:132499)
A;Note: part of this sequence was confirmed by protein sequencing
C;Keywords: membrane-associated protein

  Query Match             14.6%;  Score 70;  DB 2;  Length 755;
  Best Local Similarity   30.5%;  Pred. No. 22;
  Matches   25;  Conservative   12;  Mismatches   39;  Indels    6;  Gaps    3;

Qy         14 VTFYLILGTGTLG--HGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRI 71
              :| |       :|    |||     | : | || ||     | |:  |  :   |||  :::
Db        326 ITLYTASHGASLSPRHGGKDLTRAVQDITAAKPRLKGLPKNLPRKKEMVAKDRADEILEE 385

Qy         72 LKKRG--NMRGGEV--RESAET 89
                ||  :|: | |  :|  |:
Db        386 TLLRGPEDMKPGRVVIKEEGES 407


RESULT 7
T06692
hypothetical protein T17F15.220 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C;Accession: T06692
R;Quetier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15793
A;Accession: T06692
A;Molecule type: DNA
A;Residues: 1-1291 <QUE>
A;Cross-references: EMBL:AL049658; GSPDB:GN00061; ATSP:T17F15.220
A;Experimental source: cultivar Columbia; BAC clone T17F15
C;Genetics:
A;Gene: ATSP:T17F15.220
A;Map position: 3
A;Introns: 863/3; 906/3; 970/3; 987/2; 1091/3; 1134/1; 1201/1

  Query Match             14.6%;  Score 70;  DB 2;  Length 1291;
  Best Local Similarity   30.3%;  Pred. No. 41;
  Matches   30;  Conservative   16;  Mismatches   43;  Indels   10;  Gaps    4;

Qy          6 YLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRL 65
              |  :|:||| |:      |    | ||:  |   ||  |  |  |  ||  :|| |: |
Db        893 YQIVLLPLVKSYMRAHLEALAEKDA-TEKSDAAREALLVELALDSKKEARGRNDNSKNTL 951

Qy         66 -----TQMIRILKKRGNMR---GGEVRESAETL-HEITP 95
                  : |:  :|  :|:   | : | :|::: |   |
Db        952 EKSKDKKKIKDTRKVKDMKATIGSDHRSNADSVEHSPLP 990


RESULT 8
S33813
kinesin light chain - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C;Accession: S33813
R;Wedaman, K.P.; Knight, A.E.; Kendrick-Jones, J.; Scholey, J.M.
J. Mol. Biol. 231, 155-158, 1993
A;Title: Sequences of sea urchin kinesin light chain isoforms.
A;Reference number: S33813; MUID:93267648; PMID:8496962
A;Accession: S33813
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-649 <WED>
A;Cross-references: EMBL:L10233; NID:g161525; PID:g161526

  Query Match             14.5%;  Score 69.5;  DB 2;  Length 649;
  Best Local Similarity   31.6%;  Pred. No. 21;
  Matches   18;  Conservative   10;  Mismatches   26;  Indels    3;  Gaps    1;

Qy         21 GTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGN 77
              | | |   |:|:: |   |:  :    || |   | :||:|  :    : :|  |||
Db        565 GDGKLRRSGSLSKLR---ASVRRSSTKLLNKLKGRESDDDGGMKRASSMSVLPSRGN 618


RESULT 9
T17199
CL3BB protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T17199
R;Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A;Description: CL family.
A;Reference number: Z18712
A;Accession: T17199
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1298 <SUG>
A;Cross-references: EMBL:AF081158; NID:g3695142; PID:g3695143; PIDN:AAC62664.1
C;Superfamily: alpha-latrotoxin receptor, calcium-independent

  Query Match             14.5%;  Score 69.5;  DB 2;  Length 1298;
  Best Local Similarity   28.9%;  Pred. No. 46;
  Matches   26;  Conservative    7;  Mismatches   24;  Indels   33;  Gaps    3;

Qy          2 QLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALK----------PEPVLLQ 50
              ||:  : || |:|            |||    ||  | |  |:          |  :| |
Db          5 QLFILMMLLAPVV----------HGGKHNERHPALAAPLRHAEHSPGGPLPPRHLLQQ 52

Qy         51 KSAARSTDDNGKDRLTQMIRILKKRGNMRG 80
              :| |||    |:           ||  ||
Db         53 PAAERSTAHRGQG----------PRGTARG 72


RESULT 10
T17200
CL3BC protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T17200
R;Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A;Description: CL family.
A;Reference number: Z18712
A;Accession: T17200
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1341 <SUG>
A;Cross-references: EMBL:AF081159; NID:g3695144; PID:g3695145; PIDN:AAC62665.1
C;Superfamily: alpha-latrotoxin receptor, calcium-independent

  Query Match             14.5%;  Score 69.5;  DB 2;  Length 1341;
  Best Local Similarity   28.9%;  Pred. No. 48;
  Matches   26;  Conservative    7;  Mismatches   24;  Indels   33;  Gaps    3;

Qy          2 QLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALK----------PEPVLLQ 50
              ||:  : || |:|            |||    ||  | |  |:          |  :| |
Db          5 QLFILMMLLAPVV----------HGGKHNERHPALAAPLRHAEHSPGGPLPPRHLLQQ 52

Qy         51 KSAARSTDDNGKDRLTQMIRILKKRGNMRG 80
              :| |||    |:           ||  ||
Db         53 PAAERSTAHRGQG----------PRGTARG 72


RESULT 11
T17198
CL3BA protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T17198
```

```
R;Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A;Description: CL family.
A;Reference number: Z18712
A;Accession: T17198
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1527 <SUG>
A;Cross-references: EMBL:AF081157; NID:g3695140; PID:g3695141; PIDN:AAC62663.1
C;Superfamily: alpha-latrotoxin receptor, calcium-independent

  Query Match             14.5%;  Score 69.5;  DB 2;  Length 1527;
  Best Local Similarity   28.9%;  Pred. No. 55;
  Matches   26;  Conservative    7;  Mismatches   24;  Indels   33;  Gaps    3;

Qy          2 QLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALK----------PEPVLLQ 50
              ||:  : || |:|           |||   ||  | |  |:          |  :| |
Db          5 QLFILMMLLAPVV----------HGGKHNERHPALAAPLRHAEHSPGGPLPPRHLLQQ 52

Qy         51 KSAARSTDDNGKDRLTQMIRILKKRGNMRG 80
               :| |||   |:            ||  ||
Db         53 PAAERSTAHRGQG----------PRGTARG 72


RESULT 12
T14327
alpha-latrotoxin receptor 3, calcium-independent - rat
N;Alternate names: G-protein coupled receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C;Accession: T14327
R;Ichtchenko, K.A.; Bittner, M.A.; Krasnoperov, V.; Little, A.R.; Chepurny, O.; Holz, R.
J. Biol. Chem. 274, 5491-5498, 1999
A;Title: A novel ubiquitously expressed alpha-latrotoxin receptor is a member of the CIR
A;Reference number: Z17983; MUID:99150330; PMID:10026162
A;Accession: T14327
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1550 <ICH>
A;Cross-references: EMBL:AF063103; NID:g3882980; PID:g3882981; PIDN:AAC77816.1
C;Genetics:
A;Gene: CIRL-3
C;Superfamily: alpha-latrotoxin receptor, calcium-independent

  Query Match             14.5%;  Score 69.5;  DB 2;  Length 1550;
  Best Local Similarity   28.9%;  Pred. No. 56;
  Matches   26;  Conservative    7;  Mismatches   24;  Indels   33;  Gaps    3;

Qy          2 QLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALK----------PEPVLLQ 50
              ||:  : || |:|           |||   ||  | |  |:          |  :| |
Db          5 QLFILMMLLAPVV----------HGGKHNERHPALAAPLRHAEHSPGGPLPPRHLLQQ 52

Qy         51 KSAARSTDDNGKDRLTQMIRILKKRGNMRG 80
               :| |||   |:            ||  ||
Db         53 PAAERSTAHRGQG----------PRGTARG 72


RESULT 13
F96587
hypothetical protein T22H22.1 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: F96587
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96587
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1556 <STO>
A;Cross-references: GB:AE005173; NID:g3776579; PIDN:AAC64896.1; GSPDB:GN00141
C;Genetics:
A;Gene: T22H22.1
A;Map position: 1
C;Superfamily: myosin MYO2; myosin motor domain homology

  Query Match             14.5%;  Score 69.5;  DB 2;  Length 1556;
  Best Local Similarity   28.8%;  Pred. No. 57;
  Matches   21;  Conservative   12;  Mismatches   33;  Indels    7;  Gaps    2;

Qy         13 LVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRI- 71
               |: :   ||       | |:|| |  :   |:  |||     |::  :|   |  : : |
Db        164 LMRYLAYLG------GRAVTEGRTVEQQVLESNPVLEAFGNAKTVRNNNSSRFGKFVEIQ 217

Qy         72 LKKRGNMRGGEVR 84
                |:| : |  ||
Db        218 FDKQGRISGAAVR 230


RESULT 14
T00727
myosin heavy chain PCR43 - Arabidopsis thaliana
N;Alternate names: protein F22O13.22
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
C;Accession: T00727; S51822; S55444
R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Co
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, April 1998
A;Description: Genomic sequence for Arabidopsis thaliana BAC F22O13.
A;Reference number: Z14200
A;Accession: T00727
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1583 <SHI>
A;Cross-references: EMBL:AC003981; NID:g3063438; PID:g3063460; GSPDB:GN00059; ATSP:F22O
R;Kinkema, M.; Wang, H.; Schiefelbein, J.
Plant Mol. Biol. 26, 1139-1153, 1994
A;Title: Molecular analysis of the myosin gene family in Arabidopsis thaliana.
A;Reference number: S51820; MUID:95111097; PMID:7811972
A;Accession: S51822
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 260-263,'I',265-282,'Y',284-447 <KIN>
A;Cross-references: EMBL:Z34297
R;Kinkema, M.D.
submitted to the EMBL Data Library, June 1994
A;Reference number: S55444
A;Accession: S55444
A;Molecule type: DNA
A;Residues: 260-263,'I',265-279,'N',281-282,'Y',284-447 <KIW>
A;Cross-references: EMBL:Z34297
C;Genetics:
A;Gene: PCR43; ATSP:F22O13.22
A;Map position: 1
A;Introns: 18/1; 52/3; 100/3; 149/2; 201/3; 221/2; 274/3; 324/3; 370/2; 419/2; 453/2; 4
/3; 1265/3; 1334/3; 1382/3; 1406/2; 1439/3; 1458/3; 1477/3; 1515/3; 1543/2
C;Superfamily: myosin MYO2; myosin motor domain homology
F;73-762/Domain: myosin motor domain homology <MMOT>

  Query Match             14.3%;  Score 68.5;  DB 2;  Length 1583;
  Best Local Similarity   27.4%;  Pred. No. 74;
  Matches   20;  Conservative   13;  Mismatches   33;  Indels    7;  Gaps    2;

Qy         13 LVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRI- 71
               |: :   ||       | |:|| |  :   |:  |||     |::  :|   |  : : |
Db        177 LMRYLAYLG------GRAVTEGRTVEQQVLESNPVLEAFGNAKTVRNNNSSRFGKFVEIQ 230
```

```
Qy          72 LKKRGNMRGGEVR 84
               |:| : |  :|
Db         231 FDKQGRISGAAIR 243


RESULT 15
G69868
Hypothetical protein ykvP - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: G69868
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
r, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
F.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G69868
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-399 <KUN>
A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13251.1; PID:e1184968;
A;Experimental source: strain 168
C;Genetics:
A;Gene: ykvP

  Query Match             14.1%;  Score 67.5;  DB 2;  Length 399;
  Best Local Similarity   25.0%;  Pred. No. 20;
  Matches   28;  Conservative   14;  Mismatches   33;  Indels   37;  Gaps    4;

Qy          17 YLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARST------------------- 57
               | |||:|     |  ||     |      ||    |:  |: : |
Db         242 YEILGSG----GFLLTSDTPAVRGKFKPGRDLIVSSSPKETLEKVKYYLNHDSERKKIQI 297

Qy          58 -------DDNGKDRLTQMIRILKKRGNMRG-GE------VRESAETLHEITP 95
                    :|: : |  :|: :|| || :|  ||       | :    :| :||
Db         298 NGKKAVKNDSYRHRAERMLEVLKSRGIIRNLGETIHYVDVLKEKYVIHHVTP 349


Search completed: June  2, 2004, 18:13:09
Job time : 27.7752 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:06:18 ; Search time 17.6744 Seconds
(without alignments)
279.877 Million cell updates/sec

Title: US-10-092-367-73
Perfect score: 479
Sequence: 1 MQLYTYLYLLVPLVTFYLIL.....GNMRGGEVRESAETLHEITP 95

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	350	73.1	100	1	CXKG_CONGE	P07231	conus geogr
2	79	16.5	328	1	SCGF_RAT	O88201	rattus norv
3	70.5	14.7	686	1	KLC_STRPU	Q05090	strongyloce
4	70	14.6	593	1	APRD_PSEAE	Q03024	pseudomonas
5	70	14.6	756	1	BFS1_BOVIN	Q06002	bos taurus
6	67	14.0	789	1	CAD9_HUMAN	Q9ulb4	homo sapien
7	67	14.0	935	1	EAE_ECO11	O31000	escherichia
8	66.5	13.9	245	1	YJBG_ECOLI	P32688	escherichia
9	66.5	13.9	511	1	EX7L_BRUME	Q8yck1	brucella me
10	66.5	13.9	511	1	EX7L_BRUSU	Q8fvr1	brucella su
11	66	13.8	770	1	TOP1_THEAC	Q9hm08	thermoplasm
12	66	13.8	929	1	CA1C_NOTVI	Q91145	notophthalm
13	64.5	13.5	558	1	ATPA_MYCLE	P45825	mycobacteri
14	64.5	13.5	870	1	YKL6_CAEEL	P42173	caenorhabdi
15	64.5	13.5	1582	1	YU30_RALSO	Q8xv02	ralstonia s
16	64	13.4	157	1	DYR1_ECOLI	P00382	escherichia
17	64	13.4	556	1	HIR3_HUMAN	Q9bw71	homo sapien
18	63.5	13.3	292	1	TRUB_STRPN	Q97qj3	streptococc
19	63.5	13.3	292	1	TRUB_STRR6	Q8cwr2	streptococc
20	63.5	13.3	454	1	DAT_HAEIN	P44951	h diaminobu
21	63	13.2	267	1	PSTB_XANAC	Q8pm59	xanthomonas
22	63	13.2	267	1	PSTB_XANCP	Q8pag0	xanthomonas
23	63	13.2	520	1	YMDA_BACSU	O31774	bacillus su
24	63	13.2	549	1	YJCE_ECOLI	P32703	escherichia
25	62.5	13.0	159	1	ATPF_THIFE	P41172	thiobacillu
26	62.5	13.0	342	1	RTCA_PYRFU	Q8u0n7	pyrococcus
27	62.5	13.0	359	1	MKK2_DROME	P49071	drosophila
28	62.5	13.0	449	1	BPL3_MOUSE	Q8bu51	mus musculu
29	62.5	13.0	451	1	AGAL_SALTY	P30877	salmonella
30	62	12.9	346	1	HYPE_BRAJA	P31906	bradyrhizob
31	62	12.9	600	1	LAM2_HUMAN	Q03252	homo sapien
32	62	12.9	857	1	LOX3_SOYBN	P09186	glycine max
33	62	12.9	934	1	EAE_ECO57	P43261	escherichia
34	61.5	12.8	431	1	YMH7_CAEEL	P34474	caenorhabdi
35	61.5	12.8	504	1	SIK1_YEAST	Q12460	saccharomyc
36	61.5	12.8	4391	1	PGBM_HUMAN	P98160	homo sapien
37	61	12.7	697	1	YN26_MYCTU	P71886	mycobacteri
38	61	12.7	801	1	DHGA_ACICA	P05465	acinetobact
39	61	12.7	962	1	PTRA_ECO57	Q8x6m8	escherichia
40	61	12.7	962	1	PTRA_ECOL6	Q8cvs2	escherichia
41	61	12.7	962	1	PTRA_ECOLI	P05458	escherichia
42	61	12.7	962	1	PTRA_SHIFL	Q83qc3	shigella fl
43	60.5	12.6	183	1	OLEC_BRANA	P29526	brassica na
44	60.5	12.6	197	1	YC77_CHLTE	Q8kcy0	chlorobium
45	60.5	12.6	356	1	COL5_ARATH	Q9c9f4	arabidopsis

ALIGNMENTS

RESULT 1
CXKG_CONGE
ID CXKG_CONGE STANDARD; PRT; 100 AA.
AC P07231; O61475;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Conantokin-G precursor (Con-G) (Conotoxin GV) (Sleeper peptide).
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom duct;
RA Bandyopadhyay P.K., Colledge C.J., Walker C.S., Zhou L.M.,
RA Hillyard D.R., Olivera B.M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 81-97 FROM N.A.
RX MEDLINE=85054897; PubMed=6501296;
RA McIntosh J.M., Olivera B.M., Cruz L.J., Gray W.R.;
RT "Gamma-carboxyglutamate in a neuroactive toxin.";
RL J. Biol. Chem. 259:14343-14346(1984).
RN [3]
RP FUNCTION.
RX MEDLINE=90327072; PubMed=2165278;
RA Yoshikami D.;
RL Unpublished results, cited by:
RL Olivera B.M., Rivier J.E., Clark C., Ramilo C.A., Corpuz G.P.,
RL Abogadie F.C., Mena E.E., Woodward S.R., Hillyard D.R., Cruz L.J.;
RL Science 249:257-263(1990).
RN [4]
RP STRUCTURE BY NMR OF 81-97.
RX MEDLINE=97332451; PubMed=9188685;
RA Rigby A.C., Baleja J.D., Furie B.C., Furie B.;
RT "Three-dimensional structure of a gamma-carboxyglutamic acid-
RT containing conotoxin, conantokin G, from the marine snail Conus
RT geographus: the metal-free conformer.";
RL Biochemistry 36:6906-6914(1997).
RN [5]
RP STRUCTURE BY NMR OF 81-97.
RX MEDLINE=98062280; PubMed=9398296;
RA Rigby A.C., Baleja J.D., Li L., Pedersen L.G., Furie B.C., Furie B.;
RT "Role of gamma-carboxyglutamic acid in the calcium-induced structural
RT transition of conantokin G, a conotoxin from the marine snail Conus
RT geographus.";
RL Biochemistry 36:15677-15684(1997).
RN [6]
RP STRUCTURE BY NMR OF 81-97.
RX MEDLINE=97153002; PubMed=8999936;
RA Skjaerbaek N., Nielsen K.J., Lewis R.J., Alewood P.F., Craik D.J.;
RT "Determination of the solution structures of conantokin-G and
RT conantokin-T by CD and NMR spectroscopy.";
RL J. Biol. Chem. 272:2291-2299(1997).

```
CC   -!- FUNCTION: Induces sleep-like symptoms in young mice and
CC       hyperactivity in older mice. Inhibits N-methyl-D-aspartate (NMDA)
CC       receptor-mediated calcium influx in central nervous system
CC       neurons.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: Belongs to the conantokin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF043141; AAC15669.1; -.
DR   PIR; A05168; A05168.
DR   PDB; 1AD7; 20-AUG-97.
DR   PDB; 1ONU; 04-SEP-97.
DR   PDB; 1AWY; 08-APR-98.
KW   Toxin; Vitamin K; Gamma-carboxyglutamic acid; Calcium; Amidation;
KW   Signal; 3D-structure.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     80
FT   PEPTIDE      81     97       CONANTOKIN-G.
FT   MOD_RES      83     83       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      84     84       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      87     87       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      90     90       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      94     94       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      97     97       AMIDATION (G-98 PROVIDE AMIDE GROUP).
FT   CONFLICT     85     85       L -> V (IN REF. 1).
SQ   SEQUENCE   100 AA;  11267 MW;  3B0050FDFF2B9DFB CRC64;

  Query Match             73.1%;  Score 350;  DB 1;  Length 100;
  Best Local Similarity   78.7%;  Pred. No. 4.8e-30;
  Matches   74;  Conservative    6;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          1 MQLYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDN 60
              |:|||||||||||||:||||||| |||||||| |||||||| ||||||||||||||||
Db          1 MHLYTYLYLLVPLVTFHLILGTGTLDDGGALTERRSADATALKAEPVLLQKSAARSTDDN 60

Qy         61 GKDRLTQMIRILKKRGNMRGG--EVRESAETLHE 92
              |||||||| ||||:|||    |   |::|: | : |
Db         61 GKDRLTQMKRILKQRGNKARGEEELQENQELIRE 94


RESULT 2
SCGF_RAT
ID   SCGF_RAT       STANDARD;      PRT;   328 AA.
AC   O88201;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Stem cell growth factor precursor (Lymphocyte secreted C-type
DE   lectin).
GN   SCGF.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone;
RX   MEDLINE=98381038; PubMed=9705843;
RA   Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,
RA   Sekine S., Hiraoka A.;
RT   "Isolation and characterization of a cDNA for human, mouse, and rat
RT   full-length stem cell growth factor, a new member of C-type lectin
RT   superfamily.";
RL   Biochem. Biophys. Res. Commun. 249:124-130(1998).
CC   -!- FUNCTION: Stimulates the proliferation and differentiation of
CC       hematopoietic precursor cells from various lineages, including
CC       erythrocytes, lymphocytes, granulocytes and macrophages. Acts
CC       synergistically with other cytokines, including IL-3, GCSF, GMCSF
CC       and FLT3 ligand. Suppresses SCF-stimulated erythrocyte
CC       proliferation (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted (By
CC       similarity).
CC   -!- PTM: O-glycosylated. Probably sulfated on the O-glycans (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB009246; BAA32406.1; -.
DR   HSSP; P05452; 1HTN.
DR   GO; GO:0005576; C:extracellular; IDA.
DR   GO; GO:0008083; F:growth factor activity; IDA.
DR   GO; GO:0005529; F:sugar binding; NAS.
DR   GO; GO:0008284; P:positive regulation of cell proliferation; IDA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Growth factor; Glycoprotein; Lectin; Signal.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   CHAIN        22    328       STEM CELL GROWTH FACTOR.
FT   DOMAIN      188    325       C-TYPE LECTIN.
SQ   SEQUENCE   328 AA;  36387 MW;  C115188AC23D6257 CRC64;

  Query Match             16.5%;  Score 79;  DB 1;  Length 328;
  Best Local Similarity   34.4%;  Pred. No. 0.59;
  Matches   32;  Conservative    8;  Mismatches   27;  Indels   26;  Gaps    5;

Qy          9 LLVP-LVTFYLILGTGTLGH--------GGALTERRLADATALKPEPVLLQKSAARSTDD 59
              |||| |::|     | |  ||         |||| | |  ::  ||      ||::    |
Db         10 LLVPHLLSF----GHGARGHGKEWEGVWGGALEEERDRESLMLKN----LQEALGLPTGV 61

Qy         60 NGKDRLTQMIRILKKRGNMRGGEVRESAETLHE 92
                || | :          |  | || |: ||  |
Db         62 GNKDNLAE---------NSEGKEVWEATETQGE 85


RESULT 3
KLC_STRPU
ID   KLC_STRPU      STANDARD;      PRT;   686 AA.
AC   Q05090; Q04801; Q05088; Q05089;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Kinesin light chain (KLC).
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS KLC-1; KLC-2; KLC-3 AND KLC-4).
RC   TISSUE=Egg;
RX   MEDLINE=93267648; PubMed=8496962;
RA   Wedaman K.P., Knight A.E., Kendrick-Jones J., Scholey J.M.;
RT   "Sequences of sea urchin kinesin light chain isoforms.";
RL   J. Mol. Biol. 231:155-158(1993).
CC   -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC       PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. THE LIGHT
CC       CHAIN MAY FUNCTION IN COUPLING OF CARGO TO THE HEAVY CHAIN OR IN
CC       THE MODULATION OF ITS ATPASE ACTIVITY.
```

```
CC   -!- SUBUNIT: Oligomeric complex composed of two heavy chains and two
CC       light chains.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC         Comment=Additional isoforms seem to exist;
CC       Name=KLC-3;
CC         IsoId=Q05090-1; Sequence=Displayed;
CC       Name=KLC-1;
CC         IsoId=Q05090-2; Sequence=VSP_002878;
CC       Name=KLC-2;
CC         IsoId=Q05090-3; Sequence=VSP_002877;
CC       Name=KLC-4;
CC         IsoId=Q05090-4; Sequence=VSP_002879, VSP_002880;
CC   -!- DOMAIN: THE LIGHT CHAIN IS COMPOSED OF THREE STRUCTURAL DOMAINS: A
CC       LARGE GLOBULAR N-TERMINAL DOMAIN WHICH MAY BE INVOLVED IN BINDING
CC       TO KINESIN HEAVY CHAINS, A CENTRAL ALPHA-HELICAL COILED-COIL
CC       DOMAIN THAT MEDIATES THE LIGHT CHAIN DIMERIZATION; AND A SMALL
CC       GLOBULAR C-TERMINAL WHICH MAY PLAY A ROLE IN REGULATING
CC       MECHANOCHEMICAL ACTIVITY OR ATTACHMENT OF KINESIN TO MEMBRANE-
CC       BOUND ORGANELLES.
CC   -!- PTM: PHOSPHORYLATION MAY MODULATE THE PROCESS OF MECHANOCHEMICAL
CC       COUPLING.
CC   -!- SIMILARITY: Belongs to the kinesin light chain family.
CC   -!- SIMILARITY: Contains 6 TPR repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L10235; AAA03059.1; -.
DR   EMBL; L10234; AAA03058.1; -.
DR   EMBL; L10233; AAA03057.1; -.
DR   EMBL; L08258; AAA03060.1; -.
DR   PIR; S33813; S33813.
DR   PIR; S33814; S33814.
DR   PIR; S33815; S33815.
DR   PIR; S33816; S33816.
DR   InterPro; IPR002151; Kinesin_light.
DR   InterPro; IPR008940; Prenyl_trans.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF00515; TPR; 5.
DR   PRINTS; PR00381; KINESINLIGHT.
DR   SMART; SM00028; TPR; 5.
DR   PROSITE; PS01160; KINESIN_LIGHT; 4.
KW   Motor protein; Microtubule; Coiled coil; Repeat; TPR repeat;
KW   Alternative splicing; Phosphorylation.
FT   DOMAIN       20    160       COILED COIL.
FT   REPEAT      215    248       TPR 1.
FT   REPEAT      257    290       TPR 2.
FT   REPEAT      299    332       TPR 3.
FT   REPEAT      341    374       TPR 4.
FT   REPEAT      383    416       TPR 5.
FT   REPEAT      472    505       TPR 6.
FT   VARSPLIC    564    572       Missing (in isoform KLC-2).
FT                                /FTId=VSP_002877.
FT   VARSPLIC    564    600       Missing (in isoform KLC-1).
FT                                /FTId=VSP_002878.
FT   VARSPLIC    441    451       GKFKDNAPYGD -> VKKRKPKPAKS (in isoform
FT                                KLC-4).
FT                                /FTId=VSP_002879.
FT   VARSPLIC    452    686       Missing (in isoform KLC-4).
FT                                /FTId=VSP_002880.
SQ   SEQUENCE   686 AA;  76517 MW;  603D71186CC0364B CRC64;

  Query Match             14.7%;  Score 70.5;  DB 1;  Length 686;
  Best Local Similarity   27.8%;  Pred. No. 10;
  Matches   20;  Conservative   16;  Mismatches   31;  Indels    5;  Gaps    2;

Qy          8 YLLVPLVTFYLIL--GTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRL 65
              |: :|    ::::  | | |   |:|:: |   |:  :    || |   | :||:|  :
Db        587 YVEIPRSPPHVLVENGDGKLRRSGSLSKLR---ASVRRSSTKLLNKLKGRESDDDGGMKR 643

Qy         66 TQMIRILKKRGN 77
                : :|  |||
Db        644 ASSMSVLPSRGN 655


RESULT 4
APRD_PSEAE
ID   APRD_PSEAE     STANDARD;      PRT;   593 AA.
AC   Q03024;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alkaline protease secretion ATP-binding protein aprD.
GN   APRD OR PA1246.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=93051361; PubMed=1427098;
RA   Duong F., Lazdunski A., Cami B., Murgier M.;
RT   "Sequence of a cluster of genes controlling synthesis and secretion
RT   of alkaline protease in Pseudomonas aeruginosa: relationships to
RT   other secretory pathways.";
RL   Gene 121:47-54(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- FUNCTION: Involved in the secretion of alkaline protease.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the ABC transporter family. HlyB subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X64558; CAA45855.1; -.
DR   EMBL; AE004554; AAG04635.1; -.
DR   PIR; S26696; S26696.
DR   HSSP; P13569; 1NBD.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001140; ABC_TM_transpt.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS50929; ABC_TM1F; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   Transmembrane; Transport; ATP-binding; Complete proteome.
FT   TRANSMEM     25     45       POTENTIAL.
FT   TRANSMEM     60     80       POTENTIAL.
FT   TRANSMEM    134    154       POTENTIAL.
```

```
FT   TRANSMEM     156     176       POTENTIAL.
FT   TRANSMEM     256     276       POTENTIAL.
FT   DOMAIN       332     567       ABC TRANSPORTER.
FT   NP_BIND      366     373       ATP (BY SIMILARITY).
SQ   SEQUENCE   593 AA;  63670 MW;  CA3A817FBCC27318 CRC64;

  Query Match             14.6%;  Score 70;  DB 1;  Length 593;
  Best Local Similarity   37.3%;  Pred. No. 9.9;
  Matches   22;  Conservative    7;  Mismatches   28;  Indels    2;  Gaps    2;

Qy         19 ILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARST-DDNGKDRLTQMIRILKKRG 76
              :|| |  |  |    :|:| | ||   | |:      |  ||:|:  |   |: || ||
Db        461 VLGVGGAGLSGG-QRQRIALARALYGAPTLVVLDEPNSNLDDSGEQALLAAIQALKARG 518


RESULT 5
BFS1_BOVIN
ID   BFS1_BOVIN     STANDARD;      PRT;   756 AA.
AC   Q06002;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Filensin (Beaded filament structural protein 1).
GN   BFSP1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lens;
RX   MEDLINE=93260017; PubMed=8491777;
RA   Gounari F., Merdes A., Quinlan R., Hess J.F., Fitzgerald P.G.,
RA   Ouzounis C.A., Georgatos S.D.;
RT   "Bovine filensin possesses primary and secondary structure similarity
RT   to intermediate filament proteins.";
RL   J. Cell Biol. 121:847-853(1993).
RN   [2]
RP   REVISIONS, AND SEQUENCE FROM N.A.
RC   TISSUE=Lens;
RA   Hess J.F.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBUNIT: ASSOCIATES WITH BFSP2.
CC   -!- SUBCELLULAR LOCATION: Membrane- and cytoskeleton-associated.
CC   -!- TISSUE SPECIFICITY: Lens.
CC   -!- SIMILARITY: Belongs to the intermediate filament family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X72388; CAA51081.1; -.
DR   InterPro; IPR001664; IF.
DR   PROSITE; PS00226; IF; FALSE_NEG.
KW   Intermediate filament; Repeat; Membrane; Coiled coil; Cytoskeleton;
KW   Eye lens protein; Phosphorylation.
FT   DOMAIN         1      38       HEAD.
FT   DOMAIN        39     318       ROD.
FT   DOMAIN       319     755       TAIL.
FT   DOMAIN        39      73       COIL 1A.
FT   DOMAIN        74      82       LINKER 1.
FT   DOMAIN        83     182       COIL 1B.
FT   DOMAIN       183     199       LINKER 12.
FT   DOMAIN       200     318       COIL 2.
FT   DOMAIN       531     621       7 X 14 AA TANDEM REPEATS.
FT   REPEAT       531     544       1.
FT   REPEAT       545     551       2 (INCOMPLETE).
FT   REPEAT       552     565       3.
FT   REPEAT       566     579       4.
FT   REPEAT       580     593       5.
FT   REPEAT       594     607       6.
FT   REPEAT       608     621       7.
FT   MOD_RES        5       5       PHOSPHORYLATION (BY PKA) (POTENTIAL).
SQ   SEQUENCE   756 AA;  83001 MW;  F86A18208A8E6109 CRC64;

  Query Match             14.6%;  Score 70;  DB 1;  Length 756;
  Best Local Similarity   30.5%;  Pred. No. 13;
  Matches   25;  Conservative   12;  Mismatches   39;  Indels    6;  Gaps    3;

Qy         14 VTFYLILGTGTLG--HGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRI 71
              :| |       :|    |||    | : | || ||    | |:  |  :   |||  :::
Db        327 ITLYTASHGASLSPRHGGKDLTRAVQDITAAKPRLKGLPKNLPRKKEMVAKDRADEILEE 386

Qy         72 LKKRG--NMRGGEV--RESAET 89
                 ||  :|: | |  :|  |:
Db        387 TLLRGPEDMKPGRVVIKEEGES 408


RESULT 6
CAD9_HUMAN
ID   CAD9_HUMAN     STANDARD;      PRT;   789 AA.
AC   Q9ULB4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cadherin-9 precursor.
GN   CDH9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=10861224;
RA   Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
RT   "Identification of three human type-II classic cadherins and frequent
RT   heterophilic interactions between different subclasses of type-II
RT   classic cadherins.";
RL   Biochem. J. 349:159-167(2000).
CC   -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC       They preferentially interact with themselves in a homophilic
CC       manner in connecting cells; cadherins may thus contribute to the
CC       sorting of heterogeneous cell types.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- SIMILARITY: Contains 5 cadherin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB035302; BAA87416.1; -.
DR   HSSP; P15116; 1NCJ.
DR   Genew; HGNC:1768; CDH9.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   GO; GO:0007156; P:homophilic cell adhesion; NAS.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR000233; Cadherin_C_term.
DR   Pfam; PF00028; cadherin; 5.
DR   Pfam; PF01049; Cadherin_C_term; 1.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 5.
DR   PROSITE; PS00232; CADHERIN_1; 2.
DR   PROSITE; PS50268; CADHERIN_2; 5.
KW   Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW   Signal.
FT   SIGNAL         1      21       POTENTIAL.
FT   PROPEP        22      53       POTENTIAL.
```

```
FT   CHAIN        54    789       CADHERIN-9.
FT   DOMAIN       54    615       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    616    636       POTENTIAL.
FT   DOMAIN      637    789       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       54    159       CADHERIN 1.
FT   DOMAIN      160    268       CADHERIN 2.
FT   DOMAIN      269    383       CADHERIN 3.
FT   DOMAIN      384    486       CADHERIN 4.
FT   DOMAIN      487    608       CADHERIN 5.
FT   CARBOHYD    255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    437    437       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    455    455       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    536    536       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   789 AA;  88701 MW;  8598532756A4344F CRC64;

  Query Match             14.0%;  Score 67;  DB 1;  Length 789;
  Best Local Similarity   25.5%;  Pred. No. 28;
  Matches   25;  Conservative   13;  Mismatches   32;  Indels   28;  Gaps    4;

Qy          3 LYTYLYLLVPLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARS-----T 57
              |   | ||: :| |             ||  :|       | ||::: |   |
Db        620 LLCVLILLILVVLF------------AALKRQR-------KKEPLIISKDDVRDNIVTYN 660

Qy         58 DDNGKDRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
              |: | :  ||   |    | :|  | || ::   :: |
Db        661 DEGGGEEDTQAFDI----GTLRNPEAREDSKLRRDVMP 694


RESULT 7
EAE_ECO11
ID   EAE_ECO11      STANDARD;      PRT;   935 AA.
AC   O31000;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Intimin (Attaching and effacing protein) (Eae protein).
GN   EAE OR EAEA.
OS   Escherichia coli O111:H-.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=168927;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98187918; PubMed=9529069;
RA   Voss E., Paton A.W., Manning P.A., Paton J.C.;
RT   "Molecular analysis of Shiga toxigenic Escherichia coli O111:H-
RT   proteins which react with sera from patients with hemolytic-uremic
RT   syndrome.";
RL   Infect. Immun. 66:1467-1472(1998).
CC   -!- FUNCTION: NECESSARY FOR THE PRODUCTION OF ATTACHING AND EFFACING
CC       LESIONS ON TISSUE CULTURE CELLS. BELIEVED TO MEDIATE ADHERENCE.
CC   -!- SUBCELLULAR LOCATION: Outer surface.
CC   -!- SIMILARITY: Belongs to the intimin/invasin family.
CC   -!- SIMILARITY: Contains 1 LysM repeat.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF025311; AAC69247.1; -.
DR   InterPro; IPR003344; Big_1.
DR   InterPro; IPR003343; Big_2.
DR   InterPro; IPR003535; Intimin.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF02369; Big_1; 2.
DR   Pfam; PF02368; Big_2; 1.
DR   Pfam; PF01476; LysM; 1.
DR   PRINTS; PR01369; INTIMIN.
DR   SMART; SM00634; BID_1; 2.
DR   SMART; SM00635; BID_2; 1.
DR   SMART; SM00257; LysM; 1.
KW   Outer membrane; Virulence.
FT   REPEAT       65    113       LYSM.
SQ   SEQUENCE   935 AA;  101570 MW;  406E79CDC07DEB11 CRC64;

  Query Match             14.0%;  Score 67;  DB 1;  Length 935;
  Best Local Similarity   23.1%;  Pred. No. 34;
  Matches   24;  Conservative   13;  Mismatches   35;  Indels   32;  Gaps    3;

Qy         11 VPLVTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTD--------- 58
              || |:| :: || |||    | |:         ||    |  |::    |  |
Db        594 VP-VSFNIVSGTATLGANSATTDANGKATVTLKSSTPGQVVVSAKTAEMTSALNASAVIF 652

Qy         59 ------------------DNGKDRLTQMIRILKKRGNMRGGEV 83
                               || | :|  ::::|:   : |  |
Db        653 VEQTKASITEIKADKTTAVANGNDAVTYTVKVMKEGQPVHGHSV 696


RESULT 8
YJBG_ECOLI
ID   YJBG_ECOLI     STANDARD;      PRT;   245 AA.
AC   P32688;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein yjbG precursor.
GN   YJBG OR B4028.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94089392; PubMed=8265357;
RA   Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA   Daniels D.L.;
RT   "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT   region from 89.2 to 92.8 minutes.";
RL   Nucleic Acids Res. 21:5408-5417(1993).
CC   -!- SIMILARITY: STRONG, TO E.COLI YMCB.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00006; AAC43122.1; -.
DR   EMBL; AE000476; AAC76998.1; -.
DR   PIR; C65210; C65210.
DR   EcoGene; EG11925; yjbG.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    245       HYPOTHETICAL PROTEIN YJBG.
SQ   SEQUENCE   245 AA;  26281 MW;  57A17CE53078E972 CRC64;

  Query Match             13.9%;  Score 66.5;  DB 1;  Length 245;
  Best Local Similarity   32.1%;  Pred. No. 8.9;
  Matches   18;  Conservative    8;  Mismatches   25;  Indels    5;  Gaps    1;

Qy         29 GALTERRLADATALKPEPVLLQKSAARSTDDNGKD-----RLTQMIRILKKRGNMR 79
              ||:    || | ||: :  || : | :  | :  |      | | |: ||  |  :
Db         59 GAVISEELATAAALRQQQALLTRLAEQGADSSADDAAAINALRQQIQALKVTGRQK 114


RESULT 9
EX7L_BRUME
ID   EX7L_BRUME     STANDARD;      PRT;   511 AA.
```

```
AC   Q8YCK1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE   (Exonuclease VII large subunit).
GN   XSEA OR BMEII0527.
OS   Brucella melitensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29459;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=16M / ATCC 23456 / Biotype 1;
RX   MEDLINE=20020109; PubMed=11756688;
RA   DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA   Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA   Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA   Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA   Haselkorn R., Kyrpides N., Overbeek R.;
RT   "The genome sequence of the facultative intracellular pathogen
RT   Brucella melitensis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC   -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC       acid-insoluble oligonucleotides, which are then degraded further
CC       into small acid-soluble oligonucleotides (By similarity).
CC   -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC       or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC   -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the xseA family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE009688; AAL53769.1; ALT_INIT.
DR   PIR; AF3575; AF3575.
DR   HAMAP; MF_00378; -; 1.
DR   InterPro; IPR003753; Exonuc_VII_L.
DR   InterPro; IPR004365; tRNA_anti.
DR   Pfam; PF02601; Exonuc_VII_L; 1.
DR   Pfam; PF01336; tRNA_anti; 1.
DR   TIGRFAMs; TIGR00237; xseA; 1.
KW   Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ   SEQUENCE   511 AA;  56332 MW;  6760E9944B4600E7 CRC64;

  Query Match             13.9%;  Score 66.5;  DB 1;  Length 511;
  Best Local Similarity   34.9%;  Pred. No. 20;
  Matches   22;  Conservative    7;  Mismatches   31;  Indels    3;  Gaps    1;

Qy         31 LTERRLA---DATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGNMRGGEVRESA 87
              | ||| |    |  | |||:| :     || :    |  | :|:| :|   |  |:
Db        390 LRERRTAFTHRANRLSPEPILRRTRLTGSTLEQLDRRRDQAVRLLIERVKRRSQELDRLM 449

Qy         88 ETL 90
               ||
Db        450 RTL 452


RESULT 10
EX7L_BRUSU
ID   EX7L_BRUSU     STANDARD;      PRT;   511 AA.
AC   Q8FVR1;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE   (Exonuclease VII large subunit).
GN   XSEA OR BRA0764.
OS   Brucella suis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29461;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1330 / Biovar 1;
RX   MEDLINE=22247741; PubMed=12271122;
RA   Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA   Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA   Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA   Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA   Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA   Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT   "The Brucella suis genome reveals fundamental similarities between
RT   animal and plant pathogens and symbionts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC   -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC       acid-insoluble oligonucleotides, which are then degraded further
CC       into small acid-soluble oligonucleotides (By similarity).
CC   -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC       or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC   -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the xseA family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE014571; AAN33946.1; -.
DR   TIGR; BRA0764; -.
DR   HAMAP; MF_00378; -; 1.
DR   InterPro; IPR003753; Exonuc_VII_L.
DR   InterPro; IPR004365; tRNA_anti.
DR   Pfam; PF02601; Exonuc_VII_L; 1.
DR   Pfam; PF01336; tRNA_anti; 1.
DR   TIGRFAMs; TIGR00237; xseA; 1.
KW   Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ   SEQUENCE   511 AA;  56358 MW;  FF5B79944B4600F9 CRC64;

  Query Match             13.9%;  Score 66.5;  DB 1;  Length 511;
  Best Local Similarity   34.9%;  Pred. No. 20;
  Matches   22;  Conservative    7;  Mismatches   31;  Indels    3;  Gaps    1;

Qy         31 LTERRLA---DATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGNMRGGEVRESA 87
              | ||| |    |  | |||:| :     || :    |  | :|:| :|   |  |:
Db        390 LRERRTAFTHRANRLSPEPILRRTRLTGSTLEQLDRRRDQAVRLLIERVKRRSQELDRLM 449

Qy         88 ETL 90
               ||
Db        450 RTL 452


RESULT 11
TOP1_THEAC
ID   TOP1_THEAC     STANDARD;      PRT;   770 AA.
AC   Q9HM08;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)
DE   (Untwisting enzyme) (Swivelase).
GN   TOPA OR TA0063.
OS   Thermoplasma acidophilum.
OC   Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
```

```
OC   Thermoplasmataceae; Thermoplasma.
OX   NCBI_TaxID=2303;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 1728;
RX   MEDLINE=20479972; PubMed=11029001;
RA   Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA   Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT   "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT   acidophilum.";
RL   Nature 407:508-513(2000).
CC   -!- FUNCTION: The reaction catalyzed by topoisomerases leads to the
CC       conversion of one topological isomer of DNA to another (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC       DNA, followed by passage and rejoining.
CC   -!- MISCELLANEOUS: When a topoisomerase transiently breaks a DNA
CC       backbone bond, it simultaneously forms a protein-DNA link, in
CC       which a tyrosyl oxygen in the enzyme is joined to a DNA phosphorus
CC       at one end of the enzyme-severed DNA strand.
CC   -!- SIMILARITY: Belongs to the prokaryotic type I/III topoisomerase
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL445063; CAC11211.1; -.
DR   InterPro; IPR000380; DNA_tpisomrase.
DR   InterPro; IPR003601; DNAtopI_ATP_bind.
DR   InterPro; IPR003602; DNAtopI_DNA_bind.
DR   InterPro; IPR006171; Toprim_dom.
DR   InterPro; IPR006154; Toprim_sub.
DR   Pfam; PF01131; Topoisom_bac; 1.
DR   Pfam; PF01751; Toprim; 1.
DR   Pfam; PF01396; zf-C4_Topoisom; 1.
DR   PRINTS; PR00417; PRTPISMRASEI.
DR   SMART; SM00437; TOP1Ac; 1.
DR   SMART; SM00436; TOP1Bc; 1.
DR   SMART; SM00493; TOPRIM; 1.
DR   PROSITE; PS00396; TOPOISOMERASE_I_PROK; FALSE_NEG.
KW   Isomerase; Topoisomerase; DNA-binding; Zinc-finger; Metal-binding;
KW   Repeat; Complete proteome.
FT   ZN_FING      611     638       C4-TYPE 1.
FT   ZN_FING      673     700       C4-TYPE 2.
FT   ZN_FING      719     744       C4-TYPE 3.
FT   ACT_SITE     312     312       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   770 AA;  87667 MW;  75DA8DD7BC3B8A22 CRC64;

  Query Match             13.8%;  Score 66;  DB 1;  Length 770;
  Best Local Similarity   28.6%;  Pred. No. 35;
  Matches   24;  Conservative   15;  Mismatches   37;  Indels    8;  Gaps    3;

Qy         18 LILGTGTLGHG--GALTERRLADATALKPEPV-LLQKSAARSTDDNGKD-----RLTQMI 69
              | ||| :  |   | | ||   :   :|| |: :    | || : :  |     :| : :
Db        479 LNLGTKSTRHDIIGKLIERGFIEGNPVKPTPLGMAFIDAVRSVNSHIADPEMTAKLEEDM 538

Qy         70 RILKKRGNMRGGEVRESAETLHEI 93
                ::|    :   | || : |||:
Db        539 DRIEKNEMSKNDVVEESKKMLHEV 562


RESULT 12
CA1C_NOTVI
ID   CA1C_NOTVI     STANDARD;      PRT;   929 AA.
AC   Q91145;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Collagen alpha 1(XII) chain (Fragment).
OS   Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC   Notophthalmus.
OX   NCBI_TaxID=8316;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95246925; PubMed=7729585;
RA   Wei Y., Yang E.V., Klatt K.P., Tassava R.A.;
RT   "Monoclonal antibody MT2 identifies the urodele alpha 1 chain of type
RT   XII collagen, a developmentally regulated extracellular matrix
RT   protein in regenerating newt limbs.";
RL   Dev. Biol. 168:503-513(1995).
CC   -!- FUNCTION: Type XII collagen interacts with type I collagen-
CC       containing fibrils, the COL1 domain could be associated with the
CC       surface of the fibrils, and the COL2 and NC3 domains may be
CC       localized in the perifibrillar matrix (By similarity). Could play
CC       a developmental role in regeneration.
CC   -!- SUBUNIT: Trimer of identical chains each containing 190 kDa of
CC       nontriple-helical sequences (By similarity).
CC   -!- DEVELOPMENTAL STAGE: Expression starts at 3 days after amputation
CC       in cells of the basal layer of the wound epithelium. At day 10,
CC       expression is found in both the basal wound epithelial cells and
CC       the distal mesenchyme cells. At mid-bud and late-bud blastema
CC       stages, wound epithelium expression has decreased, whereas the
CC       mesenchyme remains strongly active in transcription and showed a
CC       tendency toward distal regionalization. Condensing cartilage shows
CC       no signal. Finally, at the late digit stage, expression becomes
CC       largely restricted to the perichondrium.
CC   -!- PTM: The triple-helical tail is stabilized by disulfide bonds at
CC       each end (By similarity).
CC   -!- PTM: Prolines at the third position of the tripeptide repeating
CC       unit (G-X-Y) are hydroxylated in some or all of the chains (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
CC       INTERRUPTED HELICES (FACIT) FAMILY.
CC   -!- SIMILARITY: Contains 2 VWFA domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U19494; AAA80217.1; -.
DR   PIR; I51027; I51027.
DR   HSSP; P02751; 1FNA.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00041; fn3; 7.
DR   Pfam; PF00092; vwa; 2.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00060; FN3; 4.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 2.
KW   Extracellular matrix; Connective tissue; Repeat; Cell adhesion;
KW   Collagen; Glycoprotein.
FT   NON_TER       1       1
FT   DOMAIN       <1      49       VWFA 1.
FT   DOMAIN       63     154       FIBRONECTIN TYPE-III 1.
FT   DOMAIN      155     245       FIBRONECTIN TYPE-III 2.
FT   DOMAIN      246     338       FIBRONECTIN TYPE-III 3.
FT   DOMAIN      339     432       FIBRONECTIN TYPE-III 4.
FT   DOMAIN      433     519       FIBRONECTIN TYPE-III 5.
FT   DOMAIN      520     612       FIBRONECTIN TYPE-III 6.
FT   DOMAIN      633     805       VWFA 2.
FT   DOMAIN      818     907       FIBRONECTIN TYPE-III 7.
FT   DOMAIN      908    >929       FIBRONECTIN TYPE-III 8.
FT   CARBOHYD    231     231       O-LINKED (XYL. . .) (CHONDROITIN SULFATE)
```

```
FT                                     (POTENTIAL).
FT   CARBOHYD     324    324          O-LINKED (XYL. . .) (CHONDROITIN SULFATE)
FT                                     (POTENTIAL).
FT   CARBOHYD     415    415          O-LINKED (XYL. . .) (CHONDROITIN SULFATE)
FT                                     (POTENTIAL).
FT   CARBOHYD      98     98          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER      929    929
SQ   SEQUENCE   929 AA;  101647 MW;  AE5D7485254FD954 CRC64;

  Query Match             13.8%;  Score 66;  DB 1;  Length 929;
  Best Local Similarity   26.7%;  Pred. No. 43;
  Matches   28;  Conservative   16;  Mismatches   43;  Indels   18;  Gaps    5;

Qy          5 TYLYLLVPLVTFYLI------LGTGTLGHGGALTERRLADATALK-PEPVLLQKSAARST 57
              | || | |   :::         | ||  :|   ||  : :   |: ||     :| | |
Db        210 TTLYNLFPDTKYHVSGVPEYQSGPGTALNGNGATEEVVGEPKNLRVSEPT--TSTAMRLT 267

Qy         58 DDNGKDRLTQMIRILKKRGNMRGGEVRE-------SAETLHEITP 95
               |    :: : :| |  |     ||:::|       |   | |: |
Db        268 WDKAPGKVQRYLRNLHSRS--AGGDIKEVTVKGDTSTTVLKELDP 310


RESULT 13
ATPA_MYCLE
ID   ATPA_MYCLE     STANDARD;      PRT;   558 AA.
AC   P45825;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase alpha chain (EC 3.6.3.14).
GN   ATPA OR ML1143.
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R., Robison K.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- FUNCTION: Produces ATP from ADP in the presence of a proton
CC       gradient across the membrane. The alpha chain is a regulatory
CC       subunit.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC       core - and CF(0) - the membrane proton channel. CF(1) has five
CC       subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC       has three main subunits: a, b and c.
CC   -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U15186; AAA63109.1; -.
DR   EMBL; AL583920; CAC31524.1; -.
DR   PIR; T09976; T09976.
DR   Leproma; ML1143; -.
DR   InterPro; IPR005294; ATP_synthF1_alph.
DR   InterPro; IPR000793; ATPase_a/b_C.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR004100; ATPase_a/bN.
DR   InterPro; IPR000790; ATPase_a_C.
DR   InterPro; IPR009005; F1_ATPase_a/bN.
DR   Pfam; PF00006; ATP-synt_ab; 1.
DR   Pfam; PF00306; ATP-synt_ab_C; 1.
DR   Pfam; PF02874; ATP-synt_ab_N; 1.
DR   ProDom; PD001099; ATPase_aC; 1.
DR   TIGRFAMs; TIGR00962; atpA; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW   Hydrolase; ATP synthesis; CF(1); ATP-binding; Hydrogen ion transport;
KW   Complete proteome.
FT   NP_BIND     172    179       ATP (BY SIMILARITY).
FT   ACT_SITE    373    373       REQUIRED FOR ACTIVITY.
SQ   SEQUENCE   558 AA;  60434 MW;  825A0B2299D28AD2 CRC64;

  Query Match             13.5%;  Score 64.5;  DB 1;  Length 558;
  Best Local Similarity   29.6%;  Pred. No. 35;
  Matches   21;  Conservative   10;  Mismatches   31;  Indels    9;  Gaps    2;

Qy         18 LILGTGTLGHGGALT-------ERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIR 70
              | ||||  ||  ::         |  | |   :  | :| :   ::   |   | ||::|:
Db        448 LFLGTG--GHLDSVPVGDVRRFETELLDHIRVAQEEILTEIRESQKLTDEAADSLTEVIK 505

Qy         71 ILKKRGNMRGG 81
              ||    ||
Db        506 SFKKGFAATGG 516


RESULT 14
YKL6_CAEEL
ID   YKL6_CAEEL     STANDARD;      PRT;   870 AA.
AC   P42173;
DT   01-NOV-1995 (Rel. 32, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein C03C10.6 in chromosome III.
GN   C03C10.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Gardner A., Berks M.;
RL   Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RA   Durbin R.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z38112; CAA86235.2; -.
DR   EMBL; Z35637; CAA86235.2; JOINED.
DR   EMBL; Z35637; CAA84690.2; -.
DR   EMBL; Z38112; CAA84690.2; JOINED.
DR   WormPep; C03C10.6; CE30720.
KW   Hypothetical protein.
SQ   SEQUENCE   870 AA;  97821 MW;  360819877643278B8 CRC64;
```

```
  Query Match             13.5%;  Score 64.5;  DB 1;  Length 870;
  Best Local Similarity   30.0%;  Pred. No. 57;
  Matches   21;  Conservative   11;  Mismatches   33;  Indels    5;  Gaps    2;

Qy          27 HGGA--LTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGNMRGGEVR 84
               | ||  : :  : :   | |||:: :|   |||      |    | ::||   :| |
Db         229 HIGAVCVMQPTIVEPEGLAPEPLIARK---RSTAVQEVRRKKSAIEVVKKSCTLRCGHCN 285

Qy          85 ESAETLHEIT 94
               |: |   | |
Db         286 EAIEMFDEET 295


RESULT 15
YU30_RALSO
ID   YU30_RALSO     STANDARD;      PRT;  1582 AA.
AC   Q8XV02;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical UPF0192 protein RSc3030 precursor.
GN   RSC3030 OR RS04727.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
CC   -!- SIMILARITY: Belongs to the UPF0192 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL646073; CAD16739.1; -.
DR   InterPro; IPR008930; Terp_cyc_toroid.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     15       POTENTIAL.
FT   CHAIN        16   1582       HYPOTHETICAL PROTEIN RSC3030.
SQ   SEQUENCE   1582 AA;  170090 MW;  8683D148F5AE3C2A CRC64;

  Query Match             13.5%;  Score 64.5;  DB 1;  Length 1582;
  Best Local Similarity   30.6%;  Pred. No. 1.1e+02;
  Matches   22;  Conservative    9;  Mismatches   22;  Indels   19;  Gaps    3;

Qy          12 PLVTFYLILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRI 71
               |||   |::| |          |  ||     | |:| : |||  |     ||   :: :
Db        1286 PLVQSLLLMGGG----------RSAAD-----PAPLLARASAAMPT----MDRAVALVWL 1326

Qy          72 LKKRGNMRGGEV 83
                 |  | ::|  |
Db        1327 QKGLGGLQGANV 1338


Search completed: June  2, 2004, 18:10:21
Job time : 19.6744 secs
```

GenCore version 5.1.6

DM protein - protein search, using sw model

Run on: June 2, 2004, 18:09:54 ; Search time 73.6434 Seconds
(without alignments)
407.018 Million cell updates/sec

Title: US-10-092-367-73
Perfect score: 479
Sequence: 1 MQLYTYLYLLVPLVTFYLIL.....GNMRGGEVRESAETLHEITP 95

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	16.1	510	16	Q8FQ82	Q8fq82 corynebacte
2	73	15.2	382	2	Q9KWV5	Q9kwv5 pseudomonas
3	73	15.2	382	2	Q849Q9	Q849q9 pseudomonas
4	72.5	15.1	299	16	Q7VT60	Q7vt60 bordetella
5	72.5	15.1	650	17	O29143	O29143 archaeoglob
6	72	15.0	935	2	Q93UI3	Q93ui3 escherichia
7	71.5	14.9	439	16	Q7UQT8	Q7uqt8 rhodopirell
8	71.5	14.9	663	17	Q97VU5	Q97vu5 sulfolobus
9	71.5	14.9	869	2	Q9EYM6	Q9eym6 escherichia
10	71.5	14.9	869	2	Q9F609	Q9f609 escherichia
11	71.5	14.9	948	2	Q9RGP3	Q9rgp3 escherichia
12	71.5	14.9	948	2	Q8RNT8	Q8rnt8 escherichia
13	71.5	14.9	948	2	Q84FQ2	Q84fq2 escherichia
14	71	14.8	304	16	Q8ZJY7	Q8zjy7 salmonella
15	70.5	14.7	299	16	Q7WC22	Q7wc22 bordetella
16	70.5	14.7	694	10	Q7XFP4	Q7xfp4 oryza sativ
17	70.5	14.7	938	2	Q8KRK8	Q8krk8 escherichia
18	70	14.6	1291	10	Q9SU54	Q9su54 arabidopsis
19	69.5	14.5	405	16	Q82MQ1	Q82mq1 streptomyce
20	69.5	14.5	1527	11	O88927	O88927 rattus norv
21	69.5	14.5	1550	11	Q9Z173	Q9z173 rattus norv
22	69.5	14.5	1556	10	Q9ZVN3	Q9zvn3 arabidopsis
23	69	14.4	1663	4	Q8WZ74	Q8wz74 homo sapien
24	68.5	14.3	299	16	Q7WQ21	Q7wq21 bordetella
25	68.5	14.3	513	4	Q8NBG0	Q8nbg0 homo sapien
26	68.5	14.3	578	11	Q7TQ53	Q7tq53 mus musculu
27	68.5	14.3	619	4	Q8NAA4	Q8naa4 homo sapien
28	68.5	14.3	723	16	Q8DDK6	Q8ddk6 vibrio vuln
29	68.5	14.3	2651	10	Q9FRR5	Q9frr5 arabidopsis
30	68	14.2	935	2	Q8VL95	Q8vl95 escherichia
31	68	14.2	935	2	Q8VL00	Q8vl00 escherichia
32	68	14.2	948	2	Q8KRL1	Q8krl1 escherichia
33	67.5	14.1	309	10	Q9AX87	Q9ax87 oryza sativ
34	67.5	14.1	399	16	O31681	O31681 bacillus su
35	67.5	14.1	524	16	Q8U7Q5	Q8u7q5 agrobacteri
36	67.5	14.1	723	16	Q87TN9	Q87tn9 vibrio para
37	67.5	14.1	1299	6	O97825	O97825 bos taurus
38	67.5	14.1	1308	6	O97828	O97828 bos taurus
39	67.5	14.1	1342	6	O97826	O97826 bos taurus
40	67.5	14.1	1351	6	O97829	O97829 bos taurus
41	67.5	14.1	1502	10	Q9SK73	Q9sk73 arabidopsis
42	67.5	14.1	1571	6	O97824	O97824 bos taurus
43	67.5	14.1	1580	6	O97827	O97827 bos taurus
44	67	14.0	170	16	Q88UM6	Q88um6 lactobacill
45	67	14.0	230	16	Q9HZN0	Q9hzn0 pseudomonas

ALIGNMENTS

RESULT 1
Q8FQ82
ID Q8FQ82 PRELIMINARY; PRT; 510 AA.
AC Q8FQ82;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transport ATP-binding protein.
GN CE1251.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP005218; BAC18061.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 510 AA; 54220 MW; EFBEDAE5754D3C38 CRC64;

Query Match 16.1%; Score 77; DB 16; Length 510;
Best Local Similarity 32.5%; Pred. No. 8.6;

```
  Matches   27;  Conservative   10;  Mismatches   24;  Indels   22;  Gaps    4;

Qy          9 LLVPLVT------FYLILGTGTLGHGGALTERRLADATALKPE--------PVL----LQ 50
              :|: |||       |  ::|| | |    |||||||| :||  :         | |
Db        160 MLIALVTLPLIPVFMWLVGTLTAG----LTERRLADLSALSDQLLDLIAGLPTLRVFRRH 215

Qy         51 KSAARSTDDNGKDRLTQMIRILK 73
              :  ||  |    :  |  : :||
Db        216 RDMAREVDRLSHNHATSTLHVLK 238


RESULT 2
Q9KWV5
ID   Q9KWV5      PRELIMINARY;      PRT;   382 AA.
AC   Q9KWV5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Periplasmic linker protein.
GN   TTGD.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=303;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20115535; PubMed=10648517;
RA   Mosqueda G., Ramos J.L.;
RT   "A set of genes encoding a second toluene efflux system in Pseudomonas
RT   putida DOT-T1 is linked to the tod genes for toluene metabolism.";
RL   J. Bacteriol. 182:937-943(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Mosqueda-cano G.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y19106; CAB72258.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008565; F:protein transporter activity; IEA.
DR   GO; GO:0009306; P:protein secretion; IEA.
DR   InterPro; IPR006143; HlyD.
DR   InterPro; IPR000437; Prok_lipoprot_S.
DR   Pfam; PF00529; HlyD; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
SQ   SEQUENCE   382 AA;  41584 MW;  EE2C751979592757 CRC64;

  Query Match             15.2%;  Score 73;  DB 2;  Length 382;
  Best Local Similarity   26.0%;  Pred. No. 16;
  Matches   25;  Conservative   21;  Mismatches   30;  Indels   20;  Gaps    4;

Qy         10 LVPLVTFYLILGTGTLGHGGALTERRLADATALKPEP---VLLQKSAARSTDDNGK---- 62
              |:||   :|::| |          ::   ::||: ||     :  ::   :||  |:
Db         12 LIPLAAIWLLVGCG---------KQETVESTAVPPEVGVYTVKAQALTLTTDLPGRTSAY 62

Qy         63 ---DRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
                 :   |   ||:||  : | ||::  | |::| |
Db         63 RVSEVRPQASGILQKRMFVEGAEVKQ-GEQLYQIDP 97


RESULT 3
Q849Q9
ID   Q849Q9      PRELIMINARY;      PRT;   382 AA.
AC   Q849Q9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Inner membrane efflux protein SepA.
GN   SEPA.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=303;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=F1;
RX   MEDLINE=97188489; PubMed=9037074;
RA   Lau P.C., Wang Y., Patel A., Labbe D., Bergeron H., Brousseau R.,
RA   Konishi Y., Rawlings M.;
RT   "A bacterial basic region leucine zipper histidine kinase regulating
RT   toluene degradation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:1453-1458(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=F1;
RA   Lau P.C.K.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U72354; AAO49780.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008565; F:protein transporter activity; IEA.
DR   GO; GO:0009306; P:protein secretion; IEA.
DR   InterPro; IPR006143; HlyD.
DR   InterPro; IPR000437; Prok_lipoprot_S.
DR   Pfam; PF00529; HlyD; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
SQ   SEQUENCE   382 AA;  41557 MW;  F0DD9A1979592757 CRC64;

  Query Match             15.2%;  Score 73;  DB 2;  Length 382;
  Best Local Similarity   26.0%;  Pred. No. 16;
  Matches   25;  Conservative   21;  Mismatches   30;  Indels   20;  Gaps    4;

Qy         10 LVPLVTFYLILGTGTLGHGGALTERRLADATALKPEP---VLLQKSAARSTDDNGK---- 62
              |:||   :|::| |          ::   ::||: ||     :  ::   :||  |:
Db         12 LIPLAAIWLLVGCG---------KQETVESTAVPPEVGVYTVKAQALTLTTDLPGRTSAY 62

Qy         63 ---DRLTQMIRILKKRGNMRGGEVRESAETLHEITP 95
                 :   |   ||:||  : | ||::  | |::| |
Db         63 RVSEVRPQASGILQKRMFVEGAEVKQ-GEQLYQIDP 97


RESULT 4
Q7VT60
ID   Q7VT60      PRELIMINARY;      PRT;   299 AA.
AC   Q7VT60;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   LysR family regulatoy protein.
GN   BP3693.
OS   Bordetella pertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=520;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX   MEDLINE=22827954; PubMed=12910271;
RA   Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA   Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA   Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA   Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA   Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA   Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA   Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA   Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA   Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA   Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT   "Comparative analysis of the genome sequences of Bordetella pertussis,
RT   Bordetella parapertussis and Bordetella bronchiseptica.";
RL   Nat. Genet. 35:32-40(2003).
DR   EMBL; BX640422; CAE43950.1; -.
KW   Complete proteome.
SQ   SEQUENCE   299 AA;  32037 MW;  3B7621048C822EE0 CRC64;

  Query Match             15.1%;  Score 72.5;  DB 16;  Length 299;
  Best Local Similarity   23.7%;  Pred. No. 14;
  Matches   23;  Conservative   18;  Mismatches   39;  Indels   17;  Gaps    2;
```

```
Qy          10 LVPLVTFYLILGTGTLGHGG-------ALTERRLADATALKPEPVLLQKSAARSTDDNGK 62
               ||  |   |   :    |::  |             :|::|   :     |:|  : :|      | |:
Db           8 LVTLALFVAVARQGSISAGARQSHLAVGAASKRISDLESALGTPLLYRTAAGVELTDAGQ 67

Qy          63 DRLTQMIRILKKRGNMRG----------GEVRESAET 89
                  |   :|:|::  :|  |          |:||  :| |
Db          68 ACLAHAVRVLQEVEHMAGVLSDFAQGVRGQVRIAANT 104
```

RESULT 5
O29143
ID O29143 PRELIMINARY; PRT; 650 AA.
AC O29143;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 3-hydroxyacyl-CoA dehydrogenase (HBD-5).
GN AF1122.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE001026; AAB90118.1; -.
DR PIR; A69390; A69390.
DR HSSP; Q16836; 1F0Y.
DR TIGR; AF1122; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006631; P:fatty acid metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006108; 3HCDH_C.
DR InterPro; IPR006176; 3HCDH_N.
DR InterPro; IPR008927; 6DGDH_C_like.
DR InterPro; IPR001753; EnCoA_hydrtse.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF00725; 3HCDH; 2.
DR Pfam; PF02737; 3HCDH_N; 1.
DR Pfam; PF00378; ECH; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 650 AA; 72453 MW; 43A075BC21EE971B CRC64;

Query Match 15.1%; Score 72.5; DB 17; Length 650;
Best Local Similarity 22.8%; Pred. No. 35;
Matches 23; Conservative 16; Mismatches 29; Indels 33; Gaps 3;

```
Qy          18 LILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGN 77
               |::| ||:|||       :|:  ||    |:|      :   |   |:   :| |:::|
Db           4 LVVGAGTMGHG-------IAEVCALAGNEVIL-CDINENILKNALSRIEWSVRKLEQKGR 55

Qy          78 MRGGE-----------------------VRESAETLHEI 93
               ::| :                      | |  |  ||:
Db          56 IKGADDVLKRLKTTTDLVEAAKEADFVIEAVVEKTEVKHEV 96
```

RESULT 6
Q93UI3
ID Q93UI3 PRELIMINARY; PRT; 935 AA.
AC Q93UI3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Intimin type gamma.
GN EAE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL559;
RA Cid D.;
RT "Association between intimin (eae) and EspB (espB) gene subtypes in
RT attaching and effacing Escherichia coli strains isolated from
RT diarrheic lambs and goat kids.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF253561; AAK48434.1; -.
DR GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0009399; P:nitrogen fixation; IEA.
DR InterPro; IPR003344; Big_1.
DR InterPro; IPR003343; Big_2.
DR InterPro; IPR008147; Gln_synt_beta.
DR InterPro; IPR003535; Intimin.
DR InterPro; IPR008964; Invasin_intimin.
DR InterPro; IPR002482; LysM.
DR Pfam; PF02369; Big_1; 2.
DR Pfam; PF02368; Big_2; 1.
DR Pfam; PF01476; LysM; 1.
DR PRINTS; PR01369; INTIMIN.
DR SMART; SM00634; BID_1; 2.
DR SMART; SM00635; BID_2; 1.
DR SMART; SM00257; LysM; 1.
DR PROSITE; PS00180; GLNA_1; 1.
SQ SEQUENCE 935 AA; 101550 MW; EFD5BD5DCE85D41A CRC64;

Query Match 15.0%; Score 72; DB 2; Length 935;
Best Local Similarity 24.0%; Pred. No. 62;
Matches 25; Conservative 13; Mismatches 34; Indels 32; Gaps 3;

```
Qy          11 VPLVTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTD--------- 58
               || |:| :: || |||   | |:          ||   |  |:|    |  |
Db         594 VP-VSFNIVSGTATLGANSATTDANGKATVTLKSSTPGQVILSAKTAEMTSALNASAVIF 652

Qy          59 -------------------DNGKDRLTQMIRILKKRGNMRGGEV 83
                                  || | :|  ::::|:   ::|  |
Db         653 VEQTKGSITEIKADKTTAVANGNDAVTYTVKVMKEGQPVQGHSV 696
```

RESULT 7
Q7UQT8
ID Q7UQT8 PRELIMINARY; PRT; 439 AA.
AC Q7UQT8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB6101.
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;

```
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294143; CAD74609.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   439 AA;  47717 MW;  5F95FB090B126F33 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 16;  Length 439;
  Best Local Similarity   33.0%;  Pred. No. 28;
  Matches   29;  Conservative   14;  Mismatches   28;  Indels   17;  Gaps    5;

Qy         20 LGTGTLGHGGALTERRLADATAL------KPEPVLLQKSAARSTDDNGKDRLTQMIRILK 73
              | :|  |    | :| | |:||:         | :  :||     : |:|   : : :|
Db        316 LDSGLSGLLTGLEQRGLLDSTAVFVTGEFGRTPKINDRSA-----EGGRDHYPRCMFMLM 370

Qy         74 KRGNMRGGEV-RESAETL----HE-ITP 95
                ||:|||:|  || :|     || |||
Db        371 AGGNVRGGQVIGESDDTASGPRHEAITP 398


RESULT 8
Q97VU5
ID   Q97VU5      PRELIMINARY;      PRT;   663 AA.
AC   Q97VU5;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   3-hydroxyacyl-CoA dehydrogenase/enoyl CoA hydratase (EC
DE   1.1.1.157).
GN   SSO2514.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR   EMBL; AE006848; AAK42645.1; -.
DR   PIR; F90423; F90423.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006631; P:fatty acid metabolism; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR006108; 3HCDH_C.
DR   InterPro; IPR006176; 3HCDH_N.
DR   InterPro; IPR008927; 6DGDH_C_like.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR000205; NAD_BS.
DR   InterPro; IPR002173; PfkB.
DR   Pfam; PF00725; 3HCDH; 2.
DR   Pfam; PF02737; 3HCDH_N; 1.
DR   Pfam; PF00378; ECH; 1.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
DR   PROSITE; PS00583; PFKB_KINASES_1; 1.
KW   Oxidoreductase; Complete proteome.
SQ   SEQUENCE   663 AA;  73038 MW;  02E45FC541DCFFF5 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 17;  Length 663;
  Best Local Similarity   26.0%;  Pred. No. 47;
  Matches   19;  Conservative   19;  Mismatches   22;  Indels   13;  Gaps    3;

Qy         18 LILGTGTLGHGGALTERRLADATALKPEPVLLQKSAARSTDDNGKDRLTQMIRILKKRGN 77
              |::| ||:|||       :|:  |:    | |    ::   :|  :|:   :  |::|
Db         10 LVVGAGTMGHG-------IAEIAAISGYKVYL-SDISQDILNNALERIRWSLSKLQER-- 59

Qy         78 MRGGEVRESAETL 90
                |:::|| :|:
Db         60 ---GQIKESVDTI 69


RESULT 9
Q9EYM6
ID   Q9EYM6      PRELIMINARY;      PRT;   869 AA.
AC   Q9EYM6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Intimin type epsilon (Fragment).
GN   EAEA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=sgo28893;
RA   Lin K.-C., Schauer D., Carville A.;
RT   "Identification of enteropathogenic E. coli in cotton-top tamarins.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF319597; AAG42487.1; -.
DR   GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   GO; GO:0009399; P:nitrogen fixation; IEA.
DR   InterPro; IPR003344; Big_1.
DR   InterPro; IPR003343; Big_2.
DR   InterPro; IPR008147; Gln_synt_beta.
DR   InterPro; IPR003535; Intimin.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF02369; Big_1; 2.
DR   Pfam; PF02368; Big_2; 1.
DR   Pfam; PF01476; LysM; 1.
DR   PRINTS; PR01369; INTIMIN.
DR   SMART; SM00634; BID_1; 2.
DR   SMART; SM00635; BID_2; 1.
DR   SMART; SM00257; LysM; 1.
DR   PROSITE; PS00180; GLNA_1; 1.
FT   NON_TER       1      1
FT   NON_TER     869    869
SQ   SEQUENCE   869 AA;  94419 MW;  51CAAE7FC5FBAA28 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 2;  Length 869;
  Best Local Similarity   20.7%;  Pred. No. 64;
  Matches   25;  Conservative   14;  Mismatches   37;  Indels   45;  Gaps    3;

Qy         14 VTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTDD---------- 59
              ||| :: || |||   | |:         ||   |  |::    |  |
Db        588 VTFSIVSGTATLGANSAKTDGNGKATVTLKSGTPGQVVVSAKTAEMTSPLNASAVTFVDQ 647

Qy         60 -----------------NGKDRLTQMIRILKKR-------------GNMRGGEVRESAE 88
                              || | :| :::::|                ||: |    :  :
Db        648 TKASITEIKADKTTAKANGSDAITYIVKVMKNNQPEANHSVTFSTNFGNLGGNSNTQIVK 707

Qy         89 T 89
              |
Db        708 T 708


RESULT 10
Q9F609
ID   Q9F609      PRELIMINARY;      PRT;   869 AA.
AC   Q9F609;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Intimin type epsilon (Fragment).
```

```
GN   EAEA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=mm7898;
RX   MEDLINE=21142643; PubMed=11230413;
RA   Mansfield K.G., Lin K.C., Newman J., Schauer D., MacKey J.,
RA   Lackner A.A., Carville A.;
RT   "Identification of Enteropathogenic Escherichia coli in Simian
RT   Immunodeficiency Virus-Infected Infant and Adult Rhesus Macaques.";
RL   J. Clin. Microbiol. 39:971-976(2001).
DR   EMBL; AF301015; AAG27704.1; -.
DR   GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   GO; GO:0009399; P:nitrogen fixation; IEA.
DR   InterPro; IPR003344; Big_1.
DR   InterPro; IPR003343; Big_2.
DR   InterPro; IPR008147; Gln_synt_beta.
DR   InterPro; IPR003535; Intimin.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF02369; Big_1; 2.
DR   Pfam; PF02368; Big_2; 1.
DR   Pfam; PF01476; LysM; 1.
DR   PRINTS; PR01369; INTIMIN.
DR   SMART; SM00634; BID_1; 2.
DR   SMART; SM00635; BID_2; 1.
DR   SMART; SM00257; LysM; 1.
DR   PROSITE; PS00180; GLNA_1; 1.
FT   NON_TER       1      1
FT   NON_TER     869    869
SQ   SEQUENCE   869 AA;  94463 MW;  FB2C9B6DEA29A046 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 2;  Length 869;
  Best Local Similarity   20.7%;  Pred. No. 64;
  Matches   25;  Conservative   14;  Mismatches   37;  Indels   45;  Gaps    3;

Qy         14 VTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTDD---------- 59
               ||| :: || |||   | |:        ||   |  |::    | |
Db        588 VTFSIVSGTATLGANSAKTDGNGKATVTLKSGTPGQVVVSAKTAEMTSPLNASAVIFVDQ 647

Qy         60 ----------------NGKDRLTQMIRILKKR-------------GNMRGGEVRESAE 88
                              || | :| :::::|               ||: |    :  :
Db        648 TKASITEIKADKTTAKANGSDAITYIVKVMKNNQPEANHSVTFSTNFGNLGGNSNTQIVK 707

Qy         89 T 89
              |
Db        708 T 708


RESULT 11
Q9RGP3
ID   Q9RGP3      PRELIMINARY;      PRT;   948 AA.
AC   Q9RGP3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Intimin type epsilon.
GN   EAE.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PMK5;
RX   MEDLINE=20072666; PubMed=10603369;
RA   Oswald E., Schmidt H., Morabito S., Karch H., Marches O., Caprioli A.;
RT   "Typing of intimin genes in human and animal enterohemorrhagic and
RT   enteropathogenic Escherichia coli: characterization of a new intimin
RT   variant.";
RL   Infect. Immun. 68:64-71(2000).
DR   EMBL; AF116899; AAF23359.1; -.
DR   GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   GO; GO:0009399; P:nitrogen fixation; IEA.
DR   InterPro; IPR003344; Big_1.
DR   InterPro; IPR003343; Big_2.
DR   InterPro; IPR008147; Gln_synt_beta.
DR   InterPro; IPR003535; Intimin.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF02369; Big_1; 2.
DR   Pfam; PF02368; Big_2; 1.
DR   Pfam; PF01476; LysM; 1.
DR   PRINTS; PR01369; INTIMIN.
DR   SMART; SM00634; BID_1; 2.
DR   SMART; SM00635; BID_2; 1.
DR   SMART; SM00257; LysM; 1.
DR   PROSITE; PS00180; GLNA_1; 1.
SQ   SEQUENCE   948 AA;  103330 MW;  7B05754019813C88 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 2;  Length 948;
  Best Local Similarity   20.7%;  Pred. No. 72;
  Matches   25;  Conservative   14;  Mismatches   37;  Indels   45;  Gaps    3;

Qy         14 VTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTDD---------- 59
               ||| :: || |||   | |:        ||   |  |::    | |
Db        596 VTFSIVSGTATLGANSAKTDGNGKATVTLKSGTPGQVVVSAKTAEMTSPLNASAVIFVDQ 655

Qy         60 ----------------NGKDRLTQMIRILKKR-------------GNMRGGEVRESAE 88
                              || | :| :::::|               ||: |    :  :
Db        656 TKASITEIKADKTTAKANGSDAITYIVKVMKNNQPEANHSVTFSTNFGNLGGNSNTQIVK 715

Qy         89 T 89
              |
Db        716 T 716


RESULT 12
Q8RNT8
ID   Q8RNT8      PRELIMINARY;      PRT;   948 AA.
AC   Q8RNT8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Intimine type epsilon.
GN   EAE.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUS-02/09/010-1;
RA   Geue L., Schnick C., Conraths F.J.;
RT   "Typing of intimin gene of a potential enterohemorrhagic Escherichia
RT   coli O165:H25 isolated from a cattle.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF479581; AAL87125.1; -.
DR   GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   GO; GO:0009399; P:nitrogen fixation; IEA.
DR   InterPro; IPR003344; Big_1.
DR   InterPro; IPR003343; Big_2.
DR   InterPro; IPR008147; Gln_synt_beta.
DR   InterPro; IPR003535; Intimin.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF02369; Big_1; 2.
```

```
DR   Pfam; PF02368; Big_2; 1.
DR   Pfam; PF01476; LysM; 2.
DR   PRINTS; PR01369; INTIMIN.
DR   SMART; SM00634; BID_1; 2.
DR   SMART; SM00635; BID_2; 2.
DR   SMART; SM00257; LysM; 1.
DR   PROSITE; PS00180; GLNA_1; 1.
SQ   SEQUENCE   948 AA;  103338 MW;  D0F6C1004632A162 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 2;  Length 948;
  Best Local Similarity   20.7%;  Pred. No. 72;
  Matches   25;  Conservative   14;  Mismatches   37;  Indels   45;  Gaps    3;

Qy         14 VTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTDD---------- 59
              ||| :: || |||   | |:        ||   |  |::    |  |
Db        596 VTFSIVSGTATLGANSAKTDGNGKATVTLKSGTPGQVVVSAKTAEMTSPLNSSAVIFVDQ 655

Qy         60 -----------------NGKDRLTQMIRILKKR-------------GNMRGGEVRESAE 88
                               || | :| :::::|                ||: |    : :
Db        656 TKASITEIKADKTTAKANGSDAITYIVKVMKNNQPEANHSVTFSTNFGNLGGNSNTQIVK 715

Qy         89 T 89
              |
Db        716 T 716


RESULT 13
Q84FQ2
ID   Q84FQ2      PRELIMINARY;      PRT;   948 AA.
AC   Q84FQ2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Intimin epsilon.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TW8023;
RX   MEDLINE=22326328; PubMed=12438362;
RA   Tarr C.L., Large T.M., Moeller C.L., Lacher D.W., Tarr P.I.,
RA   Acheson D.W., Whittam T.S.;
RT   "Molecular Characterization of a Serotype O121:H19 Clone, a Distinct
RT   Shiga Toxin-Producing Clone of Pathogenic Escherichia coli.";
RL   Infect. Immun. 70:6853-6859(2002).
DR   EMBL; AY186750; AAO27352.1; -.
DR   GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR   GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   GO; GO:0009399; P:nitrogen fixation; IEA.
DR   InterPro; IPR003344; Big_1.
DR   InterPro; IPR003343; Big_2.
DR   InterPro; IPR001434; DUF11.
DR   InterPro; IPR008147; Gln_synt_beta.
DR   InterPro; IPR003535; Intimin.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF02369; Big_1; 2.
DR   Pfam; PF02368; Big_2; 1.
DR   Pfam; PF01476; LysM; 1.
DR   PRINTS; PR01369; INTIMIN.
DR   SMART; SM00634; BID_1; 2.
DR   SMART; SM00635; BID_2; 1.
DR   SMART; SM00257; LysM; 1.
DR   TIGRFAMs; TIGR01451; B_ant_repeat; 1.
DR   PROSITE; PS00180; GLNA_1; 1.
SQ   SEQUENCE   948 AA;  103310 MW;  080195440A9401D5 CRC64;

  Query Match             14.9%;  Score 71.5;  DB 2;  Length 948;
  Best Local Similarity   20.7%;  Pred. No. 72;
  Matches   25;  Conservative   14;  Mismatches   37;  Indels   45;  Gaps    3;

Qy         14 VTFYLILGTGTLGHGGALTERRLADATALK---PEPVLLQKSAARSTDD---------- 59
              ||| :: || |||   | |:        ||   |  |::    |  |
Db        596 VTFSIVSGTATLGANSAKTDGNGKATVTLKSGTPGQVVVSAKTAEMTSPLNASAVIFVDQ 655

Qy         60 -----------------NGKDRLTQMIRILKKR-------------GNMRGGEVRESAE 88
                               || | :| :::::|                ||: |    : :
Db        656 TKASITEIKADKTTAKANGSDAITYIVKVMKNNQPEANHSVTFSTNFGNLGGNSNTQIVK 715

Qy         89 T 89
              |
Db        716 T 716


RESULT 14
Q8ZJY7
ID   Q8ZJY7      PRELIMINARY;      PRT;   304 AA.
AC   Q8ZJY7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Restriction of methylated adenine.
GN   MRR OR STM4527.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008913; AAL23345.1; -.
DR   InterPro; IPR007560; Mrr_cat.
DR   Pfam; PF04471; Mrr_cat; 1.
KW   Complete proteome.
SQ   SEQUENCE   304 AA;  33596 MW;  421DB135DDDFEF13 CRC64;

  Query Match             14.8%;  Score 71;  DB 16;  Length 304;
  Best Local Similarity   34.3%;  Pred. No. 21;
  Matches   23;  Conservative   10;  Mismatches   14;  Indels   20;  Gaps    4;

Qy         35 RLADATALKPEPVLLQ-KSAARSTDDNGKDRLTQMIRILKKRGNMRGGEVRESA-----E 88
              :|||| || |:|  ::    |||: |   ||| | :              |:|||      |
Db        116 KLADAVALDPQPESIEDDELARSSPD---DRLEQAL----------NEIRESVAEELLE 161

Qy         89 TLHEITP 95
                | :::|
Db        162 NLLQVSP 168


RESULT 15
Q7WC22
ID   Q7WC22      PRELIMINARY;      PRT;   299 AA.
AC   Q7WC22;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   LysR family regulatoy protein.
GN   BPP0506.
OS   Bordetella parapertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=519;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=12822 / ATCC BAA-587;
RX   MEDLINE=22827954; PubMed=12910271;
RA   Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA   Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA   Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA   Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA   Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA   Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA   Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA   Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA   Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA   Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT   "Comparative analysis of the genome sequences of Bordetella pertussis,
RT   Bordetella parapertussis and Bordetella bronchiseptica.";
RL   Nat. Genet. 35:32-40(2003).
DR   EMBL; BX640424; CAE36090.1; -.
KW   Complete proteome.
SQ   SEQUENCE   299 AA;  31938 MW;  7B49C30887B39B29 CRC64;

  Query Match             14.7%;  Score 70.5;  DB 16;  Length 299;
  Best Local Similarity   23.7%;  Pred. No. 23;
  Matches   23;  Conservative   18;  Mismatches   39;  Indels   17;  Gaps    2;

Qy         10 LVPLVTFYLILGTGTLGHGG-------ALTERRLADATALKPEPVLLQKSAARSTDDNGK 62
              || |  |  :   |::  |         :  :|::|       |:| : :|     | |:
Db          8 LVTLALFVAVARQGSISAGARQSHLAVGVASKRISDLENALGTPLLYRTAAGVELTDAGQ 67

Qy         63 DRLTQMIRILKKRGNMR----------GGEVRESAET 89
                 |   :|:|::  :|           ||:|| :| |
Db         68 ACLAHALRVLQEVEHMAGVLSDFAQGVGGQVRIAANT 104


Search completed: June  2, 2004, 18:12:15
Job time : 75.6434 secs
```